AD_____

Award Number: DAMD17-02-1-0631

TITLE: A Novel Strategy for Controlling the Metastatic Phenotype: Targeting the SNAG Repression Domain in the SNAIL Zinc-Finger Protein

PRINCIPAL INVESTIGATOR: Frank J. Rauscher, III, Ph.D.

CONTRACTING ORGANIZATION: The Wistar Institute
Philadelphia, PA 19104

REPORT DATE: July 2007

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-08-2007	Final	1 Apr 2002-31 Mar 2006

4. TITLE AND SUBTITLE

A Novel Strategy for Controlling the Metastatic Phenotype: Targeting the SNAG Repression Domain in the SNAIL Zinc-Finger Protein

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-02-1-0631

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Frank J. Rauscher, III, Ph.D.

E-Mail: rauscher@wistar.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The Wistar Institute
Philadelphia, PA 19104

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

BCRP Award has funded fundamental new discoveries into how a key protein involved in the metastatic process of breast cancer functions. We have discovered the following: 1. The SNAIL zinc finger repressor must bind its obligate co-repressor AJUBA, in order to repress e-cadherin and initiate metastasis. 2. The AJUBA interaction is direct and highly specific and this complex can be found at endogenous SNAIL target genes. 3. The binding requires the 20 amino acid SNAG domain of SNAIL and a single LIM domain of AJUBA. This biding forms a well-defined protein-protein interaction, which may be targeted by small molecules to interfere with the metastatic process during breast cancer progression.

15. SUBJECT TERMS
Not provided.

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	207	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

FINAL REPORT: DAMD17-02-1-0631

TITLE: A Novel Strategy for Controlling the Metastatic Phenotype: Targeting the SNAG Repression Domain in the SNAIL Zinc-Finger Protein

Statement of Work:

Task 1: Reconstitute, map and determine the specificity of the SNAIL-SNAG Domain/SNAP interaction *in vitro* and *in vivo*. (Months 1-18)

a. Isolate full-length SNAP and SNAIL cDNAs and create a panel of deletion and point mutations based on domain boundaries and subclone these constructs to bacterial and mammalian expression vectors (constitutive and inducible), which incorporate epitope tags. Purify bacterially expressed antigens, raise polyclonal antibodies, and evaluate their utility and specificity in biochemical assays of purified and cell-derived proteins. (Months 1-12)

Accomplishments:

This aim has been completely accomplished. We have obtained and/or cloned ourselves and subsequently subcloned cDNAs for SNAIL and the SNAP protein AJUBA and assembled a vast array of bacterial and mammalian expression vectors using both epitope tagged and non-epitope tagged vectors. Each of these constructs has been subject to extensive truncation and point mutagenesis in order to fine map interaction domains. Antibodies to AJUBA have been raised using purified bacterial protein using 6-histidine fusion technology. These antibodies have been affinity purified and are highly active in almost every assay we have tried including IHC, IF, Immunoprecipitation, and Chromatin Immunoprecipitation (CHIP) using protein sources from both cells and recombinant sources. These reagents have been distributed freely to the research community. We have had a bit of trouble getting these antibodies to work on paraffin embedded tissue specimens, however this is not uncommon and we continue to optimize this procedure. We have determined that they do not cross-react with other LIM domain proteins in all procedures and have utilized them as described below.

b. Perform *in vitro* interaction assays using a combination of GST chromatography, co-immunoprecipitation Perform *in vivo* interaction assays using transient co-transfection of tagged SNAP and SNAIL expression constructs and co-immunoprecipitation of cell extracts (as above). Perform hybrid *in vivo* / *in vitro* interaction assays combining cell derived (transfected or endogenous) proteins with purified protein baits in GST chromatographic and gel shift/supershift assays. Evaluate specificity of interactions using other SNAP family members. (Months 6-18)

Accomplishments:

This aim has been almost completely accomplished. Using a combination of the antibodies produced (described above) and affinity chromatography assays, we have efficiently reconstituted the SNAG domain-SNAP interaction in vitro and in vivo. Co-transfections of LEXA and GAL4 fusion to the minimal SNAG domain showed robust interaction between the AJUBA SNAP and the SNAG domain. This binding was dependent upon the most COOH terminal of the three LIM domains in AJUBA and did

not occur with other LIM domain proteins of the AJUBA family. Thus, there is tremendous specificity for interaction among these very highly related proteins The SNAG domain of SNAIL is required for AJUBA binding but there appears to be multiple contacts for protein-protein interaction among the complete 300 amino acid NH2 terminus of SNAIL. Importantly single amino acid substitutions in the SNAG domain strongly inhibited both AJUBA binding and repression implying that this interaction is critical in vivo. To confirm this we have also shown that the complex of SNAIL and AJUBA can be easily detected bound to the e-cadherin target gene using chromatin immunoprecipitation experiments. Though these experiments go way beyond the statement of work, they are the gold standard to show that a repressor-co-repressor interaction (like SNAIL-AJUBA) occurs at an endogenous target gene. We have now shown this with both engineered integrated targets and endogenous targets. Recapitulating the SNAG-SNAP interaction with completely purified proteins from bacteria has been very difficult (see below also). Simply put, both the SNAG peptide (alone) and the isolated AJUBA LIM domain are profoundly insoluble proteins. We tried multiple different strategies including co-expression, co-folding, low temperature induction etc. We modeled these studies after our highly successful expression and reconstitution of another zinc-finger repression domain, the KRAB domain. Unfortunately, these peptides behave very differently. We continue to optimize this in order to perform structural and small molecule inhibitor studies.

Statement of Work:

Task 2: Identify dominant negative and peptide-based inhibitors of the SNAG-SNAP interaction and use them as tools to manipulate SNAIL-mediated repression and E-Cadherin expression *in vivo*. (Months 8-30)

a. Generate stable cell lines with integrated metallothionein-inducible, epitope tagged SNAP in recipient cells with low E-Cadherin expression to evaluate whether E-Cadherin is reactivated by induction of SNAP expression to dominantly interfere with SNAIL repression. Clonal inducible cell lines will be assessed in for motility/invasiveness in tissue culture based assays, and for metastases in SCID mice based models. (Months 8-24)

Accomplishments:

This aim has been largely accomplished albeit with much different experimental strategies and systems than originally proposed. Initially we tried to produce inducible SNAP and SNAIL cell lines using a variety of inducible plasmid or lentiviral systems. All of them proved too leaky in expression and resulted in rapid mesenchymal differentiation making it impossible to examine clones. Instead we turned to P19 and F9 embryonal cells that express high levels of AJUBA. Using both stable and transient siRNA mediated knockdown technology, we have shown an inverse correlation between e-cadherin expression and AJUBA/SNAIL expression. Since these cell lines could not be evaluated in mice, we turned to a collaboration with Dr Gregg Longmore of Wash. U. St Louis, who was developing the xenopus egg system for looking at the role of AJUBA and SNAIL in early development. Together, our laboratories have shown that morpholino-mediated knockdown of either SNAIL or AJUBA is required for early neural crest induction and that the interaction of thee proteins is critical for that developmental pathway.

a. Create SNAG domain cell delivery peptides based on fusion of the 21 amino acid SNAG domain to TAT and/or antennapedia sequences and evaluate in assays described above for the alteration of SNAG mediated biological responses *in vivo*. (Months 12-30)

Accomplishments:
This did not work at all. We spent an extreme amount of effort to produce soluble TAT SNAG fusion peptides including nuclear localization signals and also including determinants to increase solubility and uptake by cells. The alternative strategy of chemically attaching the SNAG peptide to the penetraitin molecule also met with complete failure. We included proper controls for proteins that have been shown to be translocated to the nucleus and affect transcription processes using both these systems. We conclude that the physical characteristics of the SNAG peptide is are not conducive to this approach.

Statement of Work:
Task 3: Define the set of genes that are under control of the SNAIL-SNAP Repression, system in breast cancer cells using cDNA microarray analysis to define the complete transcriptome regulated by the SNAIL protein. (Months 18-36)

a. Using the inducible cell lines or the peptide-based inhibitor approach (above), isolate RNA populations from cells +/- induction, or +/- inhibitor treatment (Months 18-24).
b. Perform cDNA microarray analysis of genes that are differentially regulated by The Wistar Institutes Microarray facility to define the set of target genes that are repressed in SNAIL expressing cells, and identify those that are de-repressed as a result of blocking SNAIL function. . (Months 24-30)
c. Examine target gene (from 3b.) expression patterns in the inducible cell lines by Northern analysis and validate gene function by antisense mediated inhibition (Months 30-36

Accomplishments:
This work is ongoing albeit with a completely different experimental strategy. As described above in Task 2a we were unable to generate non-leaky clones of inducible SNAIL or AJUBA expressing cells. These clones were found unsuitable for Microarray analysis. Task 3 was completely dependent on this condition. However, we are now in the process of asking this important question in a different way by using our stable siRNA knockdown of AJUBA and SNAIL. We are using individual clonal knockdown cell lines to profile genes that are derepressed using microarray technology.

Additional discoveries, partially supported by this Award via experiments that were not contained in the original SOW:

1. Discovery that SNAG mediated e-cadherin repression is not mediated by post-translational modification by SUMO, unlike many other.

2. Discovery that the AJUBA protein binds directly to the PRMT5 arginine methyltranferase and places repressive marks on nucleosomal histones in SNAIL target genes.

Publications and Manuscripts in Preparation or Review/Revision supported by this Award:

1. Ayyanathan K., Lechner, M.S., Bell P, Maul G.G., Schultz D.C., Yamada Y., Tanaka K., Torigoe K., Rauscher, III F.J. 2003. Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. Genes & Dev. 17(15):1855-69. (Cover Image and see invited Commentary: Genes & Dev. 17(15): 1805-1812.

2. Venkov, C.D., Link, A.J., Jennings, J.L., Plieth, D., Inoue T., Nagai, K., Xu, C., Cimitrova, Y.N., Rauscher III, F.J., and Neilson, E.G. 2007. A proximal activator of transcription in epithelial-mesenchymal transition. J. Clin. Invest. 117(2): 482-491.

3. Peng, H., Gibson, L.C., Capili, A.D., Borden, K.L., Osborne, M. J., Harper, S.L., Speicher, D.W., Zhao, K., Marmorstein, R., Rock, T., Rauscher III, F.J. 2007. The Structurally Disordered KRAB Repression Domain is Incorporated into a Protease Resistant Core upon Binding to KAP-1-RBCC Domain. J. Mol. Biol. 370(2): 269-289.

4. Ajuba LIM proteins are Snail corepressors required for neural crest development in Xenopus. Langer, E. M., Feng, Y., Hou, Z., Rauscher III, F.J., Kroll, K.L., and Longmore, G. D. In revision. Dev. Cell.

5. PHD Domain-mediated E3 Ligase Activity Directs Intramolecular Sumoylation of an Adjacent Bromo Domain which is Required for Gene Silencing. Ivanov, A., Peng, H., Yurchenko, V., Negorev, D., Zeng, L., Fredericks, W., Maul, G., Sadofsky, M., Zhou, M.M., and Rauscher III, F.J. In revision. Mol. Cell.

6. The AJUBA LIM Domain Protein is a Co-Repressor for SNAG Domain Mediated Repression and Participates in Nucleo-Cytoplasmic Shuttling. Ayyanathan, K., Peng, H., Hou, Z., Fredericks, W.J., Goyal, R., Langer, E., Longmore, D., and Rauscher III, F.J. In preparation.

7. The LIM protein AJUBA recruits protein arginine methyltransferase 5 (PRMT5) to mediate SNAIL-dependent transcriptional repression. Hou, Z., Ivanov, A., Peng, H., Langer, W.M., Longmore, G.D., and Rauscher III, F.J. In preparation.

Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation

Kasirajan Ayyanathan, Mark S. Lechner, Peter Bell, Gerd G. Maul, David C. Schultz, Yoshihiko Yamada, Kazuhiro Tanaka, Kiyoyuki Torigoe and Frank J. Rauscher, III

Genes & Dev. 2003 17: 1855-1869; originally published online Jul 17, 2003;
Access the most recent version at doi:10.1101/gad.1102803

Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation

Kasirajan Ayyanathan, Mark S. Lechner,[1] Peter Bell, Gerd G. Maul, David C. Schultz,[2] Yoshihiko Yamada,[3] Kazuhiro Tanaka,[3] Kiyoyuki Torigoe,[3] and Frank J. Rauscher III[4]

The Wistar Institute, Philadelphia, Pennsylvania 19104-4268, USA

Heterochromatin protein 1 (HP1) is a key component of constitutive heterochromatin in *Drosophila* and is required for stable epigenetic gene silencing classically observed as position effect variegation. Less is known of the family of mammalian HP1 proteins, which may be euchromatic, targeted to expressed loci by repressor–corepressor complexes, and retained there by Lys 9-methylated histone H3 (H3-MeK9). To characterize the physical properties of euchromatic loci bound by HP1, we developed a strategy for regulated recruitment of HP1 to an expressed transgene in mammalian cells by using a synthetic, hormone-regulated KRAB repression domain. We show that its obligate corepressor, KAP1, can coordinate all the machinery required for stable gene silencing. In the presence of hormone, the transgene is rapidly silenced, spatially recruited to HP1-rich nuclear regions, assumes a compact chromatin structure, and is physically associated with KAP1, HP1, and the H3 Lys 9-specific methyltransferase, SETDB1, over a highly localized region centered around the promoter. Remarkably, silencing established by a short pulse of hormone is stably maintained for >50 population doublings in the absence of hormone in clonal-cell populations, and the silent transgenes in these clones show promoter hypermethylation. Thus, like variegation in *Drosophila*, recruitment of mammalian HP1 to a euchromatic promoter can establish a silenced state that is epigenetically heritable.

[*Keywords*: KRAB domain; KAP1; HP1; heterochromatin; gene silencing; position effect variegation]

Received April 10, 2003; revised version accepted May 29, 2003.

A recently emerging paradigm for the epigenetic control and propagation of gene expression states involves the role of chromatin structure. Though historically viewed as a passive packaging structure primarily used to assemble the enormous amount of DNA into a eukaryotic nucleus, the nucleosome with its complement of core histones has emerged as a key target for regulating gene expression (Wolffe and Hansen 2001). The dynamic regulation of chromatin organization appears to be accomplished by macromolecular protein complexes that contain enzymatic activities that modify the tails of the core histones. The constellation of these histone modifications, including acetylation, phosphorylation, ubiquitination, and methylation, create both synergistic and antagonistic signals that correlate with the transcriptional activity of a gene (Wu and Grunstein 2000). This emerging "histone code" is hypothesized to create functionally distinct subdomains in chromatin that define active versus transcriptionally silent genes (Jenuwein and Allis 2001). Histone modifications and the chromatin-associated proteins that interpret these signals may represent an epigenetic marking system responsible for setting and maintaining heritable programs of gene expression during development.

The role of histone acetylation/deacetylation [mediated by histone acetyltransferases (HATs) and histone deacetylases (HDACs), respectively] in modulating gene activity is now well established (Kuo and Allis 1998). The role of histone methylation in the regulation of chromatin structure and gene transcription has been greatly facilitated by the recent identification of histone

Present addresses: [1]Department of Bioscience and Biotechnology, Drexel University, Philadelphia, PA 19104, USA; [2]Department of Pharmacology, Case Western Reserve University, Cleveland, OH 44106-4965, USA; [3]National Institute of Dental & Craniofacial Research, National Institutes of Health, Bethesda, MD 20892-4370.
[4]Corresponding author.
E-MAIL rauscher@wistar.upenn.edu; FAX (215) 898-3929.
Article published online ahead of print. Article and publication date are at http://www.genesdev.org/cgi/doi/10.1101/gad.1102803.

methyltransferase (HMTase) enzymes (Zhang and Reinberg 2001). The discovery that the mammalian homologs of the *Drosophila melanogaster* heterochromatin protein Su(var)3-9 are H3-specific methyltransferases significantly supported the involvement of histone methylation in gene regulation (Rea et al. 2000). Further, the methylation was highly selective for Lys 9 with the methyltransferase function mapping to the evolutionarily conserved SET (SuVar3-9, Enhancer of Zeste, Trithorax) domain (Rea et al. 2000). Because histone H3, Lys 9 methylation (MeK9) is highly enriched in heterochromatin and other transcriptionally silent regions of the nucleus, it is postulated that this modification might be a vital component of the histone code that controls gene silencing.

The recent discovery that the Lys 9-methylated histone H3 (H3-MeK9) mark establishes a high affinity binding site for the heterochromatin protein 1 (HP1) family of heterochromatin proteins has provided one of the first links between a histone mark and establishment of a repressive chromatin environment for gene expression (Bannister et al. 2001; Lachner et al. 2001; Nakayama et al. 2001). The HP1 proteins are small nonhistone chromosomal proteins that are composed of an NH_2-terminal chromodomain (CD), a COOH-terminal chromoshadow domain (CSD), and a variable hinge region that separates these two domains (Wallrath 1998). The CD binds directly, with high affinity to the MeK9 residue in the histone H3 tail (Bannister et al. 2001; Lachner et al. 2001). The CSD is a homodimerization domain that directly recognizes a consensus pentapeptide sequence, PxVxL with high affinity, which is present in a growing number of nuclear proteins that may target the HP1 protein to specific genes or subnuclear compartments (Lechner et al. 2000; Smothers and Henikoff 2001).

Clues to the biological consequences of HP1–chromatin interaction have come from the study of *D. melanogaster* HP1 (Wallrath 1998). HP1 is intimately involved in the phenomenon of position effect variegation (PEV) in *D. melanogaster*. Classic PEV is observed when a transcribed euchromatic gene becomes integrated adjacent to a block of silent heterochromatin (Baker 1968; Wakimoto 1998). Transcription is silenced in only a subset of cells, and this state is stably inherited by their progeny thus leading to variegated or mosaic patterns of expression in the adult organism. The stochastic nature of PEV is hypothesized to be due to the variable spreading of heterochromatin (and thus transcriptional silencing) into adjacent regions of the genome (Locke et al. 1988). Genetic screens for modifiers of PEV led to the discovery of the suppressor allele *Su(var)2-5* that encoded HP1, a discovery consistent with its earlier identification as a heterochromatin-associated protein in polytene nuclei (James and Elgin 1986). Genetic experiments revealed that HP1 is a strong, dose-dependent modifier of PEV (Locke et al. 1988; Eissenberg et al. 1992). The HP1 protein is physically associated with both constitutive heterochromatin and adjacent variegating transgene, and likely contributes directly to the formation or stabilization of silent heterochromatin (Eissenberg et al. 1992). As would be expected, these large blocks of HP1-containing heterochromatin are highly enriched in H3-MeK9 (Nakayama et al. 2001). Thus, the discovery of the H3-MeK9–HP1 connection establishes a direct link between histone methylation and the stable epigenetic gene silencing, which occurs in or adjacent to large blocks of constitutive heterochromatin.

The above observations prompt two important questions: (1) Can HP1 also participate in the silencing of constitutively expressed euchromatic genes, that is, genes that are not physically adjacent to a large block of silent heterochromatin, and if so, are those silenced states stably heritable? (2) How are the required processes of histone deacetylation, H3-MeK9 methylation, HP1 deposition, and chromatin compaction coordinated in the nucleus? In our previous work, we have approached the latter question by characterizing the KAP1 corepressor. KAP1 displays the hallmarks of a scaffold protein that can recruit and coordinate many of the components required for HP1-mediated gene silencing to specific loci. KAP1 coordinates histone deacetylation via the recruitment of the NuRD complex (Schultz et al. 2001); histone H3 Lys 9 methylation via the action of a novel KAP1-associated histone methyltransferase named SETDB1 (Schultz et al. 2002); and direct binding and deposition of HP1 mediated by a highly conserved PxVxL motif present in KAP1 (Lechner et al. 2000) that interacts with the chromoshadow domain of HP1. Moreover, KAP1 functions as a corepressor by binding directly to the highly conserved KRAB repression domain that is present in more than 220 human zinc finger proteins, suggesting that this mechanism is likely targeted to a large number of specific loci in vivo (Peng et al. 2000; Abrink et al. 2001).

In the work presented here, we have utilized the KRAB–KAP1 system to answer the first question posed above, that is, can HP1 participate in the silencing of euchromatic genes? Toward this objective, we have created a hormone inducible system in a mammalian cell line that allows transient and reversible targeting of endogenous KAP1 and its associated activities to a highly transcribed euchromatic reporter transgene. We conclude that KAP1 coordinates the establishment of highly localized heterochromatin-like silenced states at euchromatic genes and that these states are epigenetically heritable.

Results

Hormone regulatable chimeric repressor proteins

We used a two-plasmid system composed of a regulatable chimeric repressor (Fig. 1A) and a synthetic reporter gene (Fig. 2A). The 90 amino-acid KRAB domain of Kox1 is sufficient to bind KAP1 and is a very strong, DNA-binding-dependent repressor in vivo (Margolin et al. 1994; Friedman et al. 1996). This domain [and a mutant, KRAB ($DV_{18,19}AA$), which lacks repression activity and fails to bind KAP1] was fused to the NH_2 terminus of the

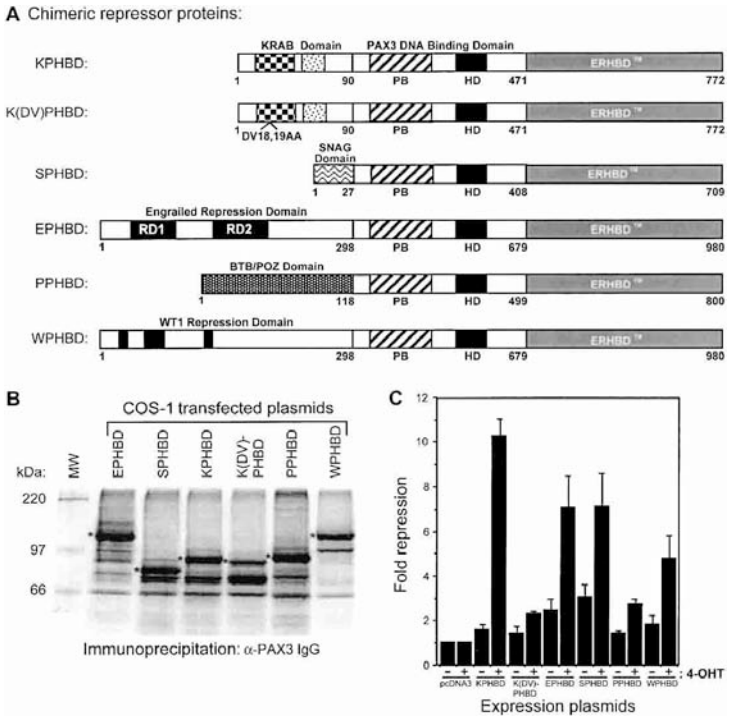


Figure 1. Characterization of chimeric transcriptional repressors. (*A*) Schematic illustration of repressor proteins. The KRAB, KRAB(DV), SNAG, Engrailed, BTB/POZ, and WT1 repression domains (RDs) were fused in frame with PAX3–HBD to generate the RD–PAX3–HBD fusion proteins. (*B*) Immunoprecipitation of transfected cell extracts with α-PAX3 IgG. Asterisks indicate the expressed proteins. MW, molecular weight markers. (*C*) 4-OHT-dependent repression of PAX3–luciferase reporter gene by the chimeric repressors. NIH3T3 cells were transfected with the indicated expression plasmids and the CD19–TK–LUC reporter plasmid. Post-transfection and 4-OHT treatment, cell lysates were assayed for luciferase and β-galactosidase activities. Fold repression represents the ratio of normalized luciferase activity of –/+OHT treated cells.

PAX3 DNA binding domain (DBD). To make this chimeric repressor hormone regulatable, we fused the estrogen receptor hormone-binding domain to the COOH terminus of the PAX3 DBD. The ERHBD contains $G_{525}R$ substitution, which renders it unresponsive to serum estrogens (Littlewood et al. 1995). These chimeras are inactive for DNA binding in the absence of hormone. For comparison purposes, we also created artificial repressors using the repression domains (RD) from WT1 (Madden et al. 1991), PLZF–BTB/POZ (Li et al. 1997), Engrailed (Jaynes and O'Farrell 1991), and GFI–SNAG (Zweidler-Mckay et al. 1996; Ayyanathan et al. 2000). Each RD–PAX3–HBD fusion protein was stably expressed in cells (Fig. 1B) and bound the PAX3 recognition sequence (data not shown). However, only the KRAB–PAX3–HBD protein could form a ternary complex with the KAP1 corepressor and the HP1 protein (data not shown). Thus, this set of chimeric repressors allows a comparison of HP1-mediated and HP1-independent mechanisms of gene silencing.

4-OHT-dependent repression of a transient luciferase reporter template

The CD19–TK–LUC–Zeo^R reporter plasmid contains six high-affinity PAX3 recognition motifs upstream of the herpes simplex virus (HSV) thymidine kinase (TK) promoter. A $zeocin^R$ gene expressed from an SV40 promoter was incorporated to provide both a selection marker allowing generation of stable cell clones and a second transcription unit linked to the luciferase gene (Ayyanathan et al. 2000). The CD19–TK–LUC–Zeo^R plasmid showed a high basal level of luciferase activity. We observed 4-OHT dependent repression of the luciferase reporter (Fig. 1C) by each RD–PAX3–HBD plasmids. The KPHBD protein was the most potent repressor (>10-fold), while the SPHBD, EPHBD, PPHBD, and WPHBD expression constructs elicited moderate levels of repression (~3- to 6-fold). The K(DV)PHBD protein was inactive for repression activity, and none of the RD–PAX3–HBD proteins repressed a TK–LUC–Zeo^R reporter, which lacked PAX3 binding sites (data not shown). We conclude that this two-plasmid system comprises a valid, hormone- and DNA-binding-dependent repression model.

4-OHT–dependent repression of chromatinized reporter transgenes

To create mammalian cell lines with integrated luciferase reporter transgenes, the DNAs for the CD19–TK–LUC–Zeo^R luciferase reporter and each of the RD–PAX3–HBD fusions were cotransfected into murine NIH3T3 fibroblasts (Fig. 2A). Following selection, clones were expanded, tested for basal luciferase activity, and then tested for 4-OHT-dependent repression activity. As controls, we also generated stable cell lines that contain only the chromatin-integrated luciferase reporter (designated as "CL" clones). At least 48 independent clonal cell lines both for CL and for each RD–PAX3–HBD fusion were tested. A representative selection of five clones for each transfection is shown in Figure 2B and C. The basal normalized luciferase activities, among clones, varied from ~10^3 to 10^6 light units/1 O.D. at A_{595} protein (Fig. 2B).

The CL clones, which do not contain a stable RD–PAX3–HBD gene, showed no response to 4-OHT. How-

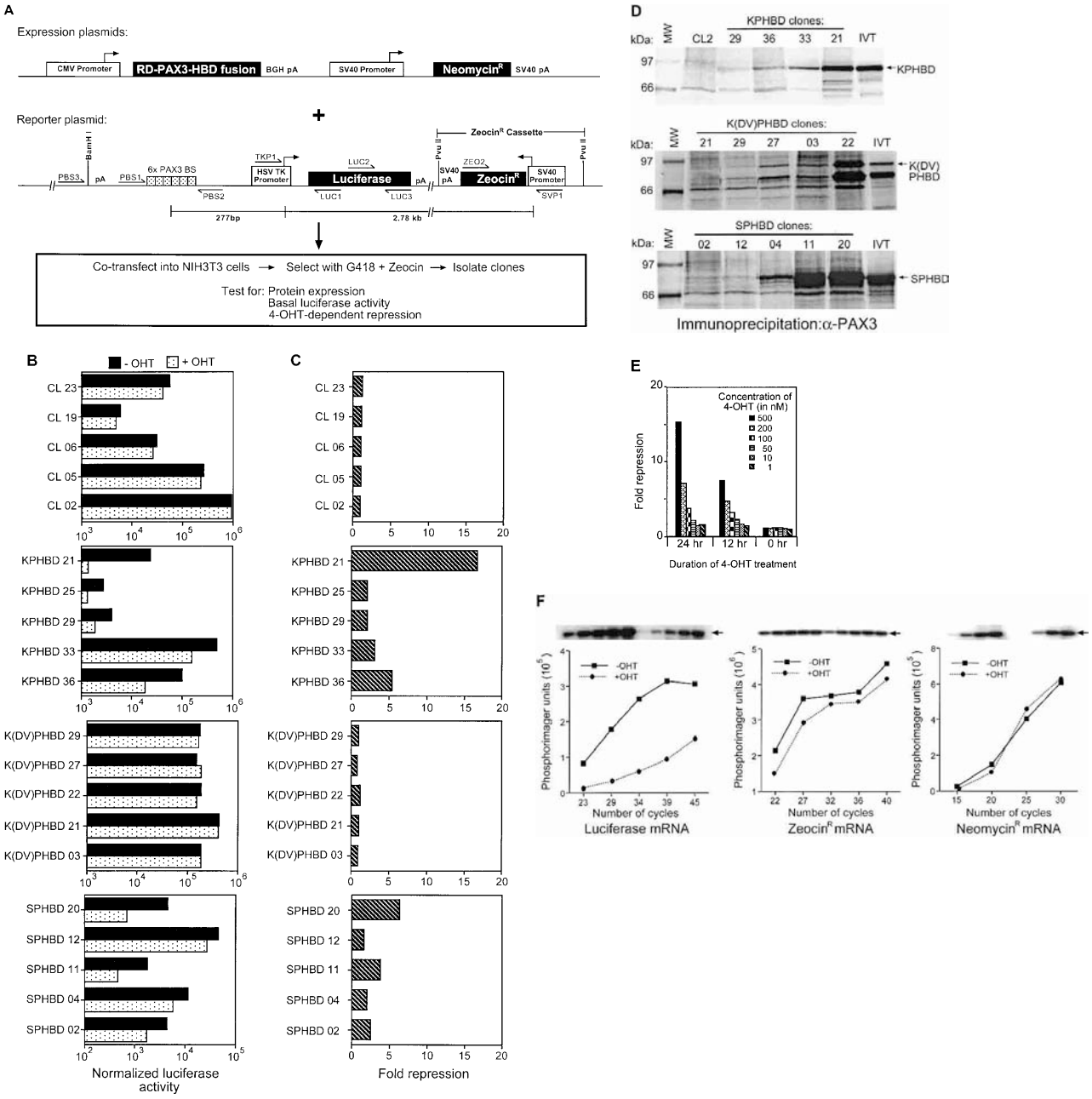


Figure 2. Characterization of stable cell lines. (*A*) Strategy to generate cell lines. (*B*,*C*) 4-OHT-dependent repression of chromatinized PAX3–luciferase reporter gene in stable cell lines. First (*top*) graphs represent the normalized luciferase activity (*B*) and fold repression (*C*) of five independent "CL" (luciferase reporter) clones. Second, third, and fourth graphs illustrate the normalized luciferase activity (*B*) and fold repression (*C*) in clones stably expressing the respective RD–PAX3–HBD proteins. (*D*) Chimeric repressor protein expression in stable cell lines. Gels illustrate the protein expression in KPHBD (*top*), K(DV)PHBD (*middle*), and SPHBD (*bottom*) stable cell lines. Expression of full-length protein is indicated by in vitro-translated protein (IVT). (*E*,*F*) Characterization of the KPHBD21 cell line. (*E*) Time- and 4-OHT-concentration-dependent repression. KPHBD21 cells were treated with vehicle or varying concentrations of 4-OHT for either 0 , 12 , or 24 h. Fold repression of luciferase activity was determined as above. (*F*) 4-OHT-dependent transcriptional repression of luciferase gene. Oligo-dT primed first strand cDNAs from vehicle or 4-OHT-treated KPHBD21 cells were monitored for luciferase, neomycin^R, and zeocin^R transcripts by quantitative PCR. Arbitrary PhosphorImager units obtained from the Southern analysis are plotted. ■, –OHT; ●, +OHT.

ever, strong 4-OHT-dependent repression was observed in KPHBD21 cell line (~16-fold) while the KPHBD30, KPHBD33, and KPHBD36 clones manifested moderate repression activity (~3- to 5-fold). In contrast, each of the K(DV)PHBD transfected clones possessed a high basal luciferase activity that was unaffected by 4-OHT treatment. The SPHBD transfection produced clones that showed six- to sevenfold repression in response to

4-OHT (Fig. 2C). Using metabolic labeling and immunoprecipitation with α-PAX3 IgG, each clone expressed full-length RD–PAX3–HBD protein to a level, which correlated roughly with the 4-OHT-dependent repression potential of these cells (Fig. 2D). Stable cell lines containing an integrated luciferase reporter and the engrailed–PAX3–HBD (EPHBD), PLZF–POZ–PAX3–HBD (PPHBD), or WT1–PAX3–HBD (WPHBD) were also generated, which showed demonstrable levels of 4-OHT-dependent repression (data not shown). We conclude the following: (1) 4-OHT treatment does not affect the basal expression of the reporter luciferase transgene when stably integrated at many different sites in the genome; (2) the KPHBD protein is a powerful, hormone-dependent repressor of the integrated reporter transgene; and (3) NIH3T3 cells contain the machinery required to support SNAG domain mediated repression.

Molecular characterization of a KRAB–PAX3–HBD stable cell line

The KPHBD21 clone was further characterized. First, a PCR using genomic DNA showed that the PAX3 binding sites, the TK promoter, the luciferase gene, and the zeocin^R cassette were physically linked in the integration site (data not shown). Second, Southern blotting suggested that an estimated two to five copies of the CD19–TK–LUC–Zeo^R were present (data not shown). Third, micrococcal nuclease digestion showed that a compact nucleosome pattern accompanied repression (data not shown). Fourth, repression was strongly time- and 4-OHT-concentration-dependent (Fig. 2E). Fifth, 4-OHT-dependent repression occurred at the level of transcription as shown by decreases in the abundance of luciferase mRNA measured by quantitative reverse transcriptase PCR (RT–PCR; Fig. 2F). However, transcription at both the linked zeocin^R locus and the unlinked neomycin^R locus were unaffected by 4-OHT treatment (Fig. 2F). Thus, KRAB-mediated repression is highly localized, as a linked transcription unit (zeocin^R cassette) ~2.8 kbp away from the repressor binding site is unaffected.

KRAB–PAX3–HBD protein induces a highly localized compact chromatin structure

To begin to dissect the changes in chromatin structure that accompany silencing, nuclease accessibility assays were performed using known restriction enzyme sites in the CD19–TK–LUC–Zeo^R plasmid. As shown in Figure 3A, each enzyme showed dramatic inhibition of cleavage in the nuclei derived from 4-OHT-treated nuclei. A region of 257 bp flanking the TK promoter and transcription initiation site in the luciferase gene is converted to a compact, nuclease-resistant structure upon 4-OHT treatment. Strong cleavage was observed at both the *Nco*I and *Stu*I sites in the presence of 4-OHT at the unlinked neomycin^R gene (Fig. 3B). Similar sites in the linked SV40 promoter–zeocin^R gene also showed equal accessibility regardless of 4-OHT treatment (Fig. 3C). Thus, the compact chromatin structure induced by KPHBD binding to the PAX3 sites near the TK promoter is not established at a promoter 2.8 kbp distant.

Recruitment of KAP1 and HP1 to the reporter luciferase transgene

We performed comprehensive chromatin immunoprecipitation (ChIP) analyses of the integrated locus and the

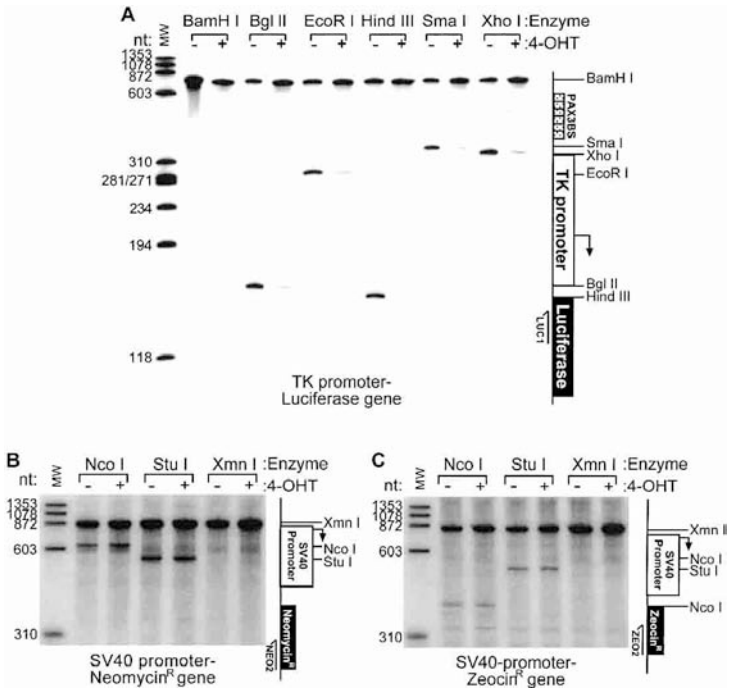


Figure 3. Differential endonuclease accessibility observed only at the luciferase locus. (*A*) Restriction endonuclease sensitivity at the TK promoter–luciferase locus. Nuclei isolated from vehicle- (–OHT) or 4-OHT (+OHT)-treated KPHBD21 cells were treated with the indicated restriction endonucleases. Purified DNA was used in reiterative primer-extension PCR with LUC1 primer. The denatured products were resolved on Urea-PAGE and autoradiographed. (*B*,*C*) Restriction endonuclease sensitivity at the neomycin^R (*B*) and zeocin^R (*C*) loci monitored by using NEO2 and ZEO2 primers. MW, radiolabeled ϕX174 *Hae*III digest; nt, nucleotide length; PAX3BS, PAX3-binding site.

results are presented in Figure 4. Using antibodies specific to the PAX3, KAP1, HP1α and HP1γ proteins, the DNA recovered was analyzed by quantitative PCR using primer pairs for regions of the CD19–TK–LUC, zeocin loci (Fig. 2A). Fragments, which bracket the PAX3 binding sites, were strongly enriched (5–10-fold) in the PAX3 immunoprecipitates (IPs) after 4-OHT-treatment. Other components of the KRAB repression complex (i.e., KAP1, HP1α, HP1γ) were inducibly recruited to the target gene. This enrichment was most evident for the primer pair that directly flanks the PAX3 binding sites (PBS1 and PBS2). This fragment is enriched 10-, 11-, and ~5-fold in PAX3, KAP1, and HP1α IPs, respectively. No signal (NS) was detected from the CL2 reporter cell line, which lacks a KPHBD repressor expression. The 257 bp fragment spanning the TK promoter and transcription initiation site is most highly enriched in KAP1 and HP1α IPs, 10- and 27-fold, respectively. This is different from HP1γ, which is constitutively bound to that region and not enriched upon 4-OHT-treatment. No signal was obtained for the 3′ end of the luciferase-coding region (located 1.2 kbp downstream from the TK promoter; primers LUC2 and LUC3), the SV40-zeocin^R cassette, or the unlinked SV40-neomycin^R cassette (data not shown). Thus, recruitment of KAP1 and HP1α is highly specific for a target gene regulated by a DNA-bound KRAB repression domain. Moreover, the association of KAP1 and HP1α to the chromatin occurs in a highly localized region.

Recruitment of KAP1 and HP1 to an endogenous target gene

To determine whether similar physical characteristics occur at an endogenous target gene that is silenced by a naturally occurring KRAB zinc finger protein, we studied the NT2 KRAB–ZFP (Fig. 5A). NT2 protein is abundant in NIH3T3 cells (Fig. 5B), binds to a 24-bp consensus sequence in the *Col11a2* promoter, and silences it (Tanaka et al. 2002; Fig. 5C). NT2 is highly transcribed while the *Col11a2* transcript is undetectable suggesting that NT2 is contributing to *Col11a2* silencing in NIH3T3 cells. However, the tightly linked *RXR*β gene, which is present 5′ of *Col11a2* is highly expressed suggesting that NT2 mediated silencing is not long range (Fig. 5C,D). Moreover, ChIP experiments demonstrated NT2 at the predicted binding site but not enriched in the proximal promoter region. However KAP1, HP1α, HP1γ, and SETDB1 proteins were abundantly present at both regions. The promoter region was also highly enriched in H3-MeK9 (Fig. 5E). These results were strikingly similar to those observed with the integrated transgene and strongly suggest that a localized heterochromatin structure is generated at an endogenous locus silenced by a KRAB–ZFP.

Spatial relocalization of the luciferase transgene to condensed chromatin

We performed fluorescence in situ hybridization (FISH) analysis on the KPHBD21 cell line using a probe for the CD19–TK–LUC luciferase reporter plasmid, and the cells were counter-stained with either DAPI or Hoechst dyes to identify condensed heterochromatin structures (Fig. 6A,B). A single locus was observed in the KPHBD21 line in every cell examined, strongly suggesting a single integration site for the reporter plasmid (data not shown). In the absence of 4-OHT, the FISH signals were most frequently found to be spatially distinct from the

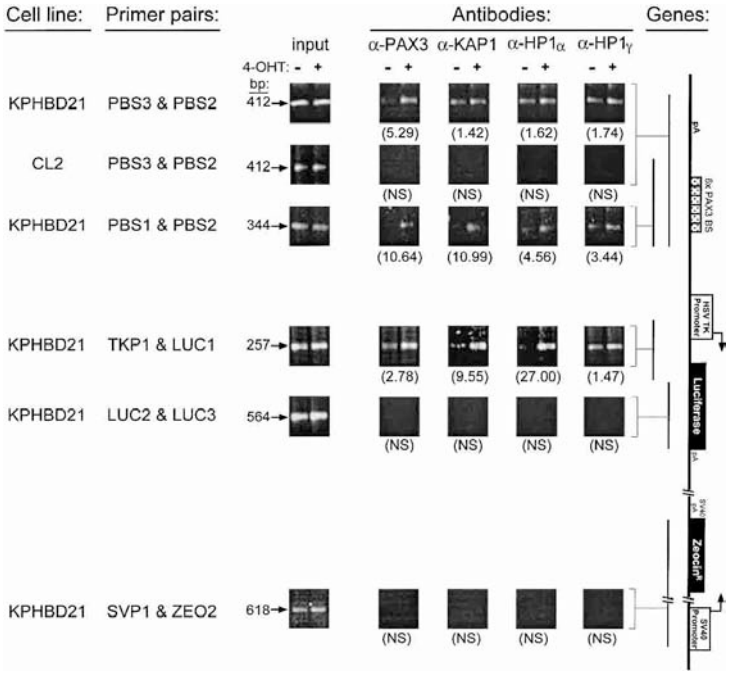


Figure 4. KRAB-box-dependent recruitment of KAP1 and HP1 proteins to the integrated luciferase transgene. CL2 and KPHBD21 cells were treated with either vehicle (–OHT) or 4-OHT (+OHT), and the DNA–protein complexes were chemically cross-linked in vivo with formaldehyde. Soluble, sonicated chromatin was immunoprecipitated with α-PAX3, α-KAP1, α-HP1α, and α-HP1γ antibodies. Quantitative PCR was carried out with input and immunoprecipitated DNAs using specific primer pairs indicated in Figure 2A. PCR-amplified DNA fragments are depicted from ChIPs with α-PAX3, α-KAP1, α-HP1α, and α-HP1γ IgG. Bold vertical lines on the *right* indicate the relative positions of the transgene fragments amplified. The numbers in parentheses indicate fold enrichment, and NS denotes no signal.

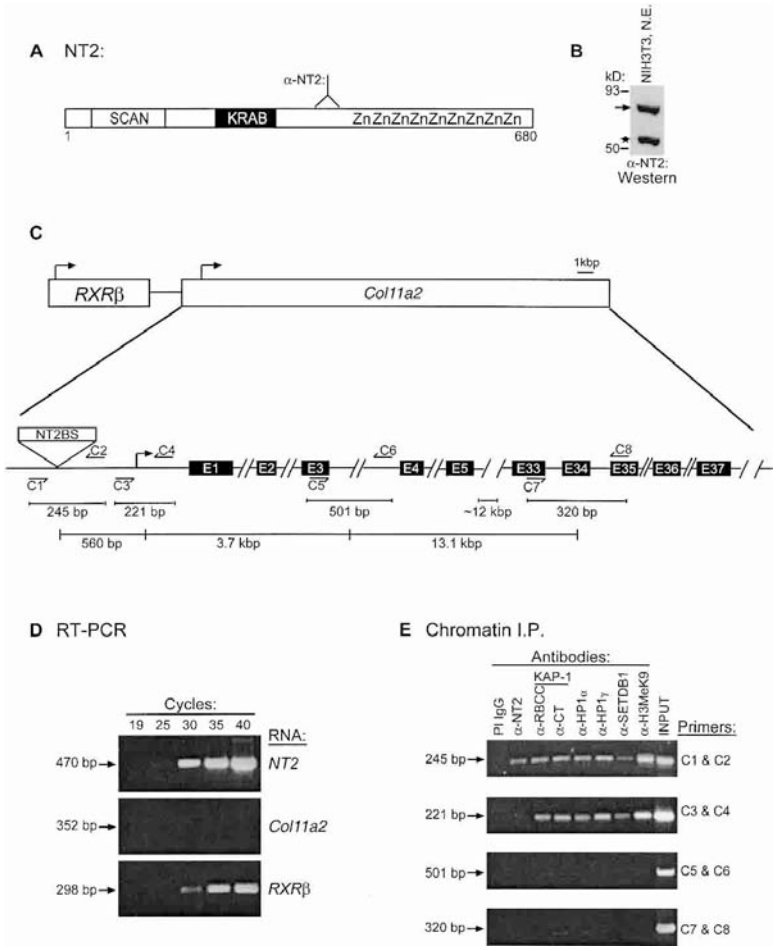


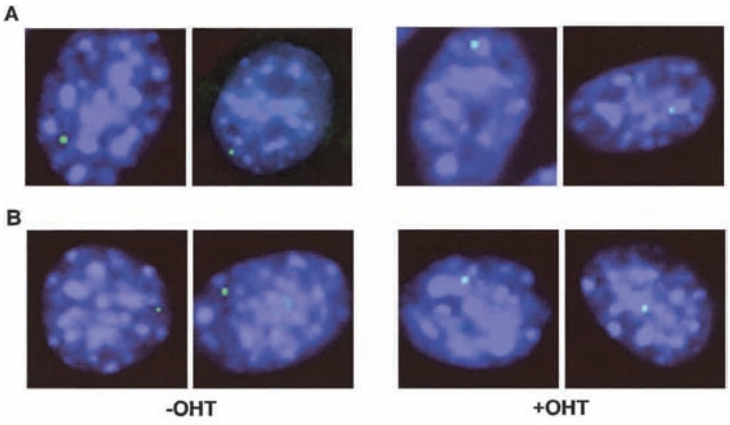
Figure 5. NT2–KRAB zinc finger protein stably represses the expression of endogenous *Col11a2* gene in NIH3T3 cells. (*A*) Diagrammatic representation of the mouse NT2–KRAB zinc finger protein. The positions of SCAN, KRAB domains, ZF motifs, and the antigen used for α-NT2 antibody production are indicated. (*B*) Abundant expression of NT2 protein in NIH3T3 cells. Nuclear proteins (100 μg) were electrophoresed on a 10% SDS-PAGE and immunoblotted with affinity-purified α-NT2 antibody. The ~75 kD NT2 protein is indicated by an arrow and a nonspecific band is marked by an asterisk. (*C*) Genomic structure of *RXR*β and *Col11a2* genes. The *Col11a2* gene is organized in a head-to-tail orientation with *RXR*β. The relative locations of oligonucleotides used in ChIP-PCR experiments and the fragment sizes are indicated. E1 to E37 represent the exons of *Col11a2* gene. (*D*) Expression patterns of *NT2*, *Col11a2*, and *RXR*β transcripts in NIH3T3 cells. Oligo-dT-primed first-strand cDNAs were amplified by quantitative PCR using primer pairs specific for the *NT2*, *Col11a2*, and *RXR*β transcripts and electrophoresed on 1.5% agarose gel. (*E*) Components of KRAB repression machinery are enriched at the silenced *Col11a2* locus. Cross-linked chromatin was immunoprecipitated with the indicated antibodies and the bound DNA was analyzed by PCR using primers indicated in *C*.

condensed chromatin islands as single green dots. However, following addition of 4-OHT, there was a clear association of the majority of luciferase FISH signals with the condensed chromatin territories. We quantitated this pattern by analyzing many cells (Fig. 6C). Using HP1α antibodies, which stained the heterochromatic islands detected by the DNA dyes, 4-OHT-dependent spatial recruitment of the luciferase FISH signals to the HP1-rich regions was also observed (data not shown). Thus, spatial relocalization of the reporter gene appears to accompany KRAB–KAP1–HP1-dependent silencing in this system.

Transient exposure of the luciferase reporter transgene to KRAB–PAX3–HBD induces stable silencing

A hallmark property of HP1-dependent gene silencing is that the silenced state is stably maintained through many cell divisions (Lu et al. 1998). To determine if there is a stable component to KRAB–KAP1–HP1-mediated gene silencing in the KPHBD21 cell line, we performed 4-OHT washout experiments. We took advantage of the fact that 4-OHT is readily removed from cells and that its effects on HBD fusion proteins are readily reversible (Littlewood et al. 1995; Pelengaris et al. 1999). A 24-h, 4-OHT treatment produced ~24-fold reduction in luciferase activity (Fig. 7A; filled bars). Twenty-four hours after hormone removal, the luciferase activity was only two- to threefold lower than a duplicate dish, which had not been treated with 4-OHT, thus showing substantial reversal of the repressed state and verifying that the hormone was removed. However, dishes harvested at 2, 3, and 4 d after 4-OHT removal still maintained a substantially lower basal luciferase level than an untreated duplicate dish. The hormone had no effect on cell growth rate and the cells went through about four rounds of cell division during the 4-d time period. This result was even more pronounced when the initial 4-OHT treatment was for 48 h (Fig. 7A, patterned bars): Dishes maintained in the absence of hormone for 4 d showed approximately sixfold lower basal luciferase activity compared to untreated dishes. This result suggested that gene silencing might be stably maintained in a subpopulation of cells.

To further explore this phenomenon, we performed longer-term, 4-OHT treatment and washout studies (Fig. 7B). A 2-, 4-, or 6-d 4-OHT treatment of KPHBD21 cells strongly repressed the reporter luciferase activity (>50-fold reduction in luciferase activity at 6 d). However, following 4-OHT removal, activity was substantially, but never completely, recovered. This effect was still evident at 8 d post-4-OHT removal, a time course that



C

Experiment:	4-OHT:	No. of cells analyzed:	% of FISH signals:	
			Associated with counter-stain	NOT associated with counter-stain
Luciferase-FISH with DAPI counter-stain	-	90	27	73
Luciferase-FISH with DAPI counter-stain	+	156	69	31
Luciferase-FISH with Hoechst counter-stain	-	82	29	71
Luciferase-FISH with Hoechst counter-stain	+	188	72	28

Figure 6. 4-OHT-dependent association of the luciferase transgene with condensed chromatin regions. The KPHBD21 cells were seeded onto glass coverslips and treated with either vehicle (–OHT) or 4-OHT (+OHT). FISH was carried out using luciferase probe and the nuclear DNA was visualized by staining with either DAPI (*A*) or Hoechst (*B*). The number of cells analyzed for each condition is indicated (*C*). The results are presented as the percent of FISH signals that were either associated (whitish green dots) or not associated (green dots) with the DAPI- or Hoechst-stained spots.

included extensive washing of each dish every day, and included a trypsinization and replating at semiconfluent cell densities. No repression was observed in two cell clones containing KRAB(DV)–PAX3–HBD fusion protein (Fig. 7C). Moreover, though substantial repression was observed in each of two SNAG–PAX3–HBD cell clones after a 4-d hormone treatment, the luciferase activity was fully recovered following 4-OHT removal (Fig. 7D).

KRAB–KAP1–HP1 mediated gene silencing is mitotically heritable over many cell generations

HP1-dependent gene variegation reflects the ability to maintain gene silencing over many cell generations in a clone of cells. To determine if a similar effect occurred at a euchromatic gene bound by HP1, we performed luciferase activities in cell populations that were transiently pulsed with 4-OHT. Three subclones of the KPHBD21 cell line, KPHBD21-08, KPHBD21-39, and KPHBD21-49, were further studied. Duplicate dishes of cells were treated with 4-OHT or vehicle for 4 d, then washed, and subjected to a limiting dilution cloning. Clones were expanded in the absence of 4-OHT for ~40 population doublings and tested for basal luciferase activity (Fig. 8A). For KPHBD21-08, 69 untreated subclones were isolated: their normalized luciferase activities showed a mean of ~10^5 light units/O.D. of protein and varied <7-fold from lowest to highest. However, subline KPHBD21-08 cells that had received 4-OHT prior to single cell cloning yielded a set of subclones with a dramatic skewing of activities. More than one-third of 65 clones yielded basal luciferase activities substantially lower than the lowest subclone derived from the untreated population (Fig. 8B). Remarkably, a portion of the clones showed barely detectable luciferase activities.

A statistical analysis using Fisher's test (F-test), which compares the variances of two samples, yielded a highly significant F value of 0.00000003. Identical results were obtained with subclones KPHBD21-39 and KPHBD21-49, where the F values comparing luciferase activity in untreated and treated populations were F = 0.00074 and F = 0.00014, respectively (Fig. 8B). The subclones maintain resistance to zeocin and G418 and thus have not deleted the transgenes. Moreover, the stably repressed clones have maintained the silent state for >10 mo in culture (data not shown). These results strongly suggest that transient recruitment of the KRAB–KAP1–HP1 complex to a euchromatic gene produces a silenced state that is mitotically heritable and does not require a persistent DNA-binding component.

Molecular characteristics of the target locus in the silenced and expressed clones

To understand the molecular basis for this stable silencing, we characterized the physical state of the silent and expressed luciferase genes. We selected silent (cl 39–40) and expressed (cl 39–45) subclones whose normalized luciferase activities were three orders of magnitude different (Fig. 9A, top panel). Using ChIP, we observed a strong enrichment of KAP1 (11.66-fold), SETDB1 (10.33-fold), HP1α (6.93-fold), H3-MeK9 (5.3-fold), and a moderate enrichment of HP1γ (2.43-fold) proteins in the silent clone compared to the expressed clone (Fig. 9A, bottom panel). However, the level of the KRAB–PAX3–HBD component at the transgene was very low and similar in the silent

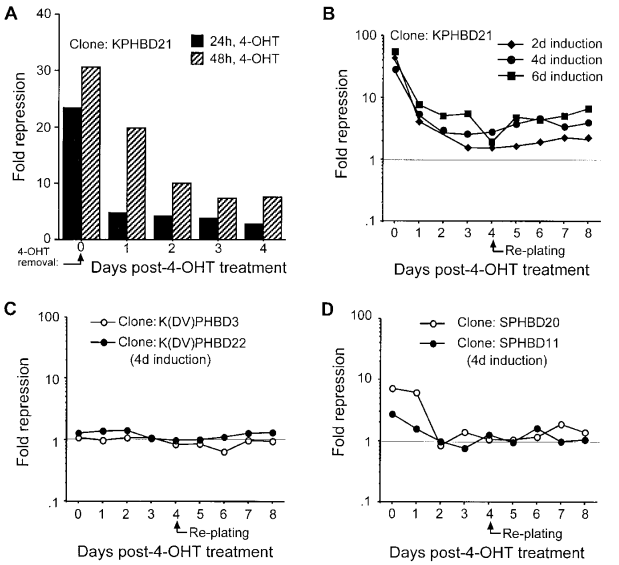


Figure 7. The KPHBD protein induces stable repression of the luciferase transgene. (*A*) Transient 4-OHT treatment of KPHBD21 cells induces a stable silencing component. Duplicate dishes of KPHBD21 cells were treated with either vehicle or 4-OHT (+OHT) continuously for 24 or 48 h and washings performed. Pairs of dishes (–/+OHT) were harvested at the indicated time points post 4-OHT removal and fold repression was determined. Filled bars, 24-h 4-OHT treatment; patterned bars, 48-h 4-OHT treatment. (*B*) A longer 4-OHT treatment of KPHBD21 cells can lead to a larger stable silencing component. Duplicate dishes of KPHBD21 cells were treated with either vehicle or 4-OHT continuously for 2, 4, or 6 d. Washings and replatings were performed as described in Materials and Methods. Pairs of dishes (–/+OHT) were harvested every day and fold repression was determined. ◆, 2 d of 4-OHT treatment; ●, 4 d of 4-OHT treatment; ■, 6 d of 4-OHT treatment. (*C*) 4-OHT-induced repression is not observed in K(DV)PHBD clones. The K(DV)PHBD3 (○) and K(DV)PHBD22 (●) clonal cell lines were treated with vehicle or 4-OHT continuously for 4 d, washed extensively, and fold repression determined. (*D*) 4-OHT-induced silencing component is not observed in SPHBD clones. The SPHBD11 (●) and SPHBD20 (○) clonal cell lines were treated with vehicle or 4-OHT continuously for 4 d, washed extensively, and fold repression determined.

and active clones. To verify that the PAX3 DNA binding sites are present and accessible in the silent clone, we transfected a VP16–PAX3 plasmid that functions as a powerful activator. We observed a dose-dependent activation of the silent luciferase locus (Fig. 9B). These results suggest that the molecular components of the KRAB–KAP1–HP1 repression pathway are constitutively present at the silenced locus and also that the maintenance of the silent state is not the result of persistent DNA binding of the KRAB–PAX3–HBD protein.

To further characterize the silent clone, we treated the cells with 5-azacytidine (5AC) and trichostatin-A (TSA). Treatment with either 5AC or TSA alone did not significantly reactivate the locus (Fig. 9C). However, sequential treatment with 5AC followed by TSA was highly synergistic in reactivating the silent locus (Fig. 9C) suggesting that the silent state may be maintained as a result of DNA methylation (Cameron et al. 1999). To confirm this, we determined the methylation status of the TK promoter regions of the silent and active clones by methylation-specific PCR (MSPCR; Fig. 9D; Herman et al. 1996). The methylation-specific primer set (M) generated a strong amplification product using DNA from the silent clone. However, no product was seen using DNA from the active clone. Conversely, the unmethylated specific primers (U) only produced a product from the active clone. These results suggest that the silent and active clones have different CpG methylation profiles at the promoter.

To map regions of methylation, we sequenced the cloned PCR products derived from bisulfite-treated genomic DNAs (Fig. 9E). The active clone showed nonrandom CpG methylation that was highly restricted to the region immediately surrounding the transcription start site. However, the silent clone showed dramatically en-

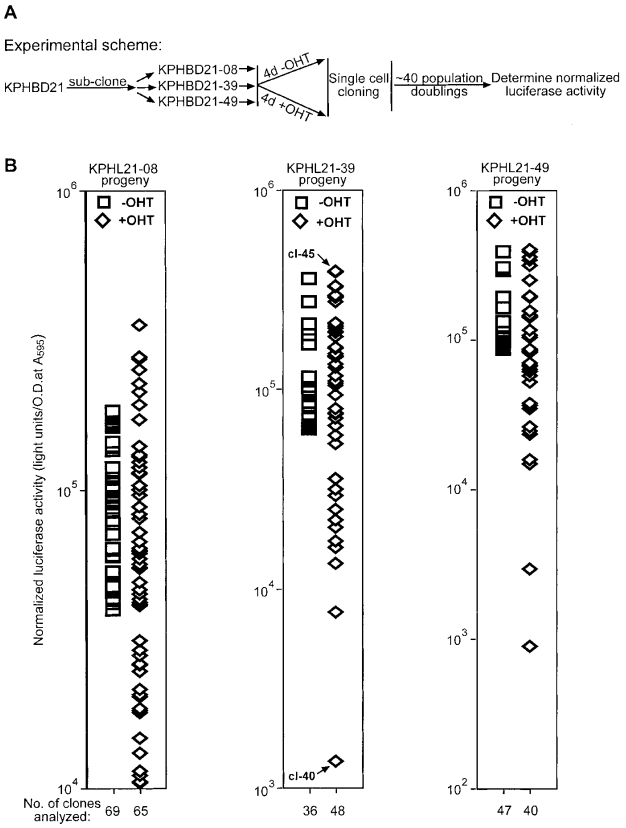


Figure 8. Clonal analysis of luciferase gene expression in KPHBD21 cells. (*A*) An experimental scheme designed to measure variegated expression of the luciferase gene in KPHBD21 cells. (*B*) Normalized luciferase activities for single cell progeny derived from KPHBD21-8, KPHBD21-39, or KPHBD21-49. Each □ and ◇ represents an independent single-cell subclone of the parental line treated with either vehicle (–OHT) or 4-OHT (+OHT), respectively. The total number of single-cell subclones analyzed for each parental line is indicated at the *bottom*. Arrows denote the silent (cl 39–40) and active (cl 39–45) clones studied in detail.

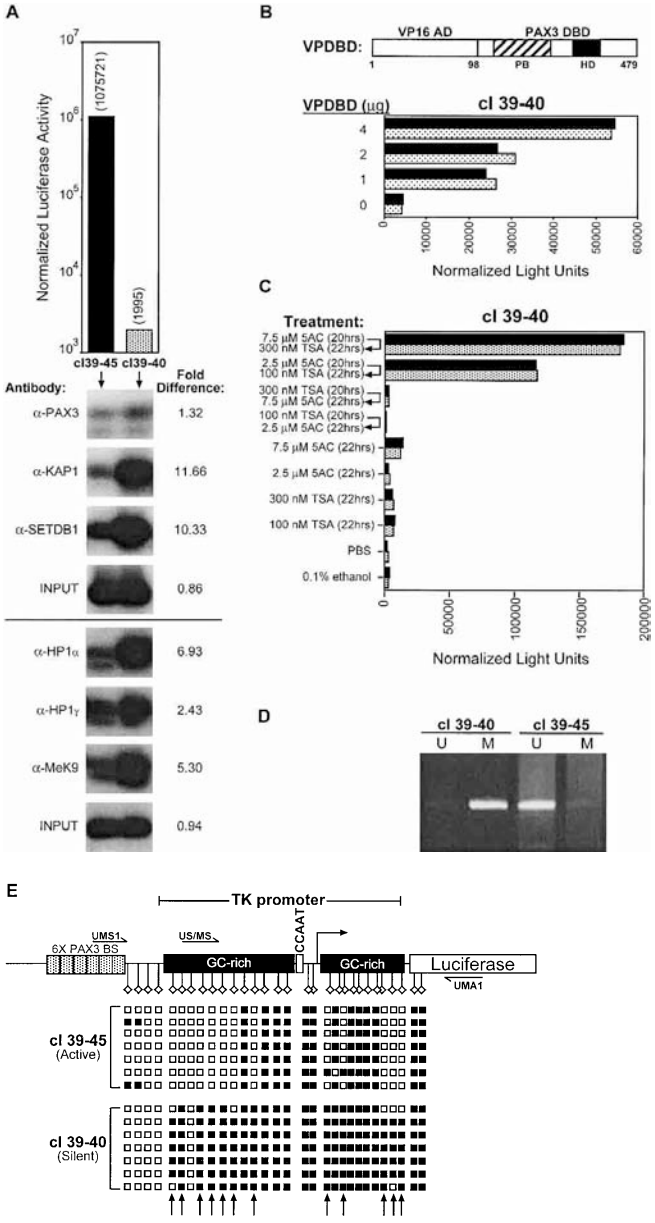


Figure 9. Molecular characteristics of the active and silent clones. (*A*) Prior to ChIP experiments, the silent (cl 39–40) and active (cl 39–45) clones were tested for their basal luciferase activities (*top* panel). Next, the cross-linked chromatin fractions obtained from silent (cl 39–40) and active (cl 39–45) clones were used in ChIPs with anti-PAX3, KAP1, SETDB1, HP1α, HP1γ, and MeK9 antibodies. The input and the immunoprecipitated DNAs were amplified using TKP1 and LUC1 primers (Fig. 2A). Fold difference was determined by PhosphorImager analysis of the Southern blots (*bottom* panel). (*B*) Indicated concentrations of VPDBD plasmid was transiently transfected into the silent clone (cl 39–40), the cell lysates were assayed for luciferase activities and normalized with β-galactosidase values. The filled and patterned bars represent results obtained from two independent experiments. (*C*) The silent (cl 39–40) and active (cl 39–45) clones were plated in multiple dishes and treated with the indicated chemicals for specified durations. The filled and patterned bars represent the luciferase activity obtained from duplicate dishes. (*D*) MSPCR. Genomic DNAs extracted from the silent (cl 39–40) and active (cl 39–45) clones were treated with sodium bisulfite and amplified using unmethylated sense (US) or methylated sense (MS) primers and an antisense primer (UMA1) belonging to the TK promoter indicated in *E* and shown. (*E*) Sodium bisulfite-genomic sequencing. The bisulfite-modified genomic DNAs of the silent (cl 39–40) and active (cl 39–45) clones were PCR-amplified using UMS1 (sense) and UMA1 (antisense) primers. The PCR products were TA-cloned into pCR II vector (Invitrogen). A representative selection of clones from both cl 39–40 and cl 39–45 were sequenced and the results presented. Hanging ◇, the positions of CpG residues; □, the unmethylated CpG; ■, methylated CpG residues. Arrows pointing up mark the CpG residues that are preferentially methylated in the genomic clones of cl 39–40.

hanced density of CpG methylation that was spread both 5′ and 3′ of the transcription start site. A total of 12 new CpG sites were methylated in the silenced clone (Fig. 9E). Together, these results suggest that DNA methylation may contribute to the mitotically heritable gene silencing we observe.

Discussion

Until recently, most of our insights into HP1 protein function and the role of heterochromatin in gene regulation have come from the study of position effect variegation (PEV) in *Drosophila*. Some of the key principles established include the following: pericentromeric and telomeric heterochromatin is enriched in DNA repeats that are packaged in a compact, chromatin structure. These regions have a low gene density, are late-replicating, and appear highly suppressed for both recombination and transcription (Wallrath 1998). The large blocks of constitutive heterochromatin are highly enriched in the HP1 protein, which can function as exquisite dose-sensitive silencers of an adjacent euchromatic transgene. Once established, this silencing is mitotically heritable through many cell divisions giving rise to the classic variegated gene expression phenotype in the adult organs. Whether these principles will hold true in higher metazoans is a key question.

Recent evidence suggests a broader role for HP1 and heterochromatin in mammalian gene regulation. First, the human and mouse genomes contain at least three HP1 homologs, some of which display constitutive euchromatic localization (Ryan et al. 1999; Minc et al.

2000). Second, heterodimerization can occur among these HP1 isoforms (Nielsen et al. 2001). Third, there are a host of regulatory proteins, which target HP1 to specific loci (Lechner et al. 2000; Smothers and Henikoff 2001). Finally, there is distinct tissue-, cell-type specificity of the HP1 isoforms including an emerging role in tumor progression (Kirschmann et al. 2000).

In this article, we have established some of the first principles for the mechanism of action of mammalian HP1 at a euchromatic locus. The KRAB–KAP1 repression system we used is a physiologically relevant targeting mechanism for HP1 proteins and, in addition, coordinates two other activities that are likely to be required for silencing: recruitment of the NuRD–HDAC complex (Schultz et al. 2001) and recruitment of the methyltransferase, SETDB1 (Schultz et al. 2002). Our results support the following conclusions: (1) The KRAB domain is able to coordinate machinery for strong transcriptional repression of an integrated, chromatinized RNA pol II transcribed target gene; (2) repression is relatively short-range; (3) repression is accompanied by a highly localized chromatin compaction; (4) the KAP1 corepressor, and HP1α/γ proteins are physically associated with the repressed gene in a highly localized manner; (5) the induced, silent state of the transgenic reporter is mitotically heritable in the absence of hormone for at least 40 population doublings; and (6), stable silencing is apparently specific for the KRAB–KAP1–HP1 mechanism as other repression domain fusions (which do not bind HP1) do not induce a heritable, silenced state. Together, these results strongly suggest that we have established a model for HP1-dependent silencing and variegation of a euchromatic gene in a mammalian cell line.

A critical issue for interpretation of these results is the chromatin environment of the integrated reporter transgene. We strongly favor the designation of the luciferase reporter as euchromatic based upon the following functional criteria: (1) All clones show a high basal level of luciferase activity; (2) all clones both express, and show physical linkage to the zeocin resistant cassette; (3) the reporter genes are packaged into a nuclease-accessible chromatin structure; and (4) interphase FISH shows the expressed reporter transgene present in euchromatic chromosome territories. Thus, while designation of a gene as euchromatic is historically based upon cytological analyses, clearly, the reporter plasmids have integrated into a region permissive for a high level of expression.

The KRAB box functions as a short range repressor

A clear result with this system is that the mechanism of KRAB-mediated repression appears to function over relatively short distances in chromatin. This property of a repression domain, that is, the ability to mediate long-range versus short-range repression has emerged as a key determinant of the biological function for a repressor protein (Gray and Levine 1996). KRAB-mediated, short-range repression was accompanied by highly localized chromatin compaction and a physical association of the KPHBD protein at the PAX3 DNA recognition sequences. The KAP1 and HP1 proteins were enriched at the TK promoter. Remarkably, the physical association of these proteins appears to only span a few nucleosomes because a region ~1.2 kbp away was devoid of cross-linkable protein. Such a short-range repressor property is also exhibited by the retinoblastoma (Rb) repressor protein, which has been shown to remodel a single nucleosome at the promoter region (Morrison et al. 2002). In light of these results, it is remarkable that we also detect spatial relocalization of the gene in the interphase nucleus to a block of constitutive heterochromatin. Intuitively, one would think that such a long-range relocalization adjacent to a large block of A-T-rich condensed chromatin would have long-range silencing effects on the gene. While one caveat to this result is the potential presence of a cryptic insulator/boundary element between the luciferase and zeocin genes, we favor the notion that silencing mediated by a short-range repressor may occur by highly localized looping into heterochromatic environments (Seum et al. 2001).

KRAB–KAP1–HP1-mediated gene silencing is mitotically heritable

The most striking finding in this study is that transgene silencing is apparently mitotically heritable in cell culture. This is the hallmark property that distinguishes HP1-dependent PEV from other mechanisms of repression. Instead of physical linkage via translocation of the euchromatic gene to adjacent heterochromatin as occurs in PEV, we have relied on transient targeting of HP1 to a transcribed locus using the KRAB–KAP1 system. Our results suggest that a pulse of KRAB–KAP–HP1 protein induces a stable, silenced state that can be detected at high frequency in clonal subpopulations following growth of >40 cell generations in the absence of hormone. Because the biological effects of HBD fusion proteins are readily reversible following removal of hormone, we conclude that the stable silencing we observe is being maintained in the absence of the KPHBD DNA binding activity. This is further supported by the observation that SNAG domain-mediated repression (which does not involve HP1) is completely reversible by hormone withdrawal under the same experimental conditions.

A key observation is that silencing did not appear to spread along the template into the adjacent zeocin transcription unit. Spreading is commonly observed in pericentromeric heterochromatin-mediated PEV in *Drosophila*. Because we continually selected for zeocin expression (via drug selection) during growth of the single-cell clones after the 4-OHT-pulse, this selection pressure may have both impeded spreading and the establishment of a larger domain of silencing. Therefore, we repeated the clonal analysis in the absence of zeocin to determine if a higher frequency of stably silenced clones could be observed. However, the same frequency of stably silenced clones was observed in the absence of zeocin when compared to the +zeocin experiment (data not

shown). We conclude that silencing of euchromatic genes via localized recruitment of HP1 may be unique to KRAB repressor domain and also that KRAB is fundamentally different from the other repressor domains in this property.

HP1 mediated silencing of euchromatic genes

Our study is one of the first to document the physical characteristics of a variegating euchromatic gene repressed by HP1 in a mammalian cell. Previously, a murine CD2 transgene reporter was shown to variegate in the thymus in response to HP1 dose (Festenstein et al. 1999). Interestingly, variegation could be either enhanced or repressed by HP1 dose, depending upon integration site (Festenstein et al. 1999). Recently, a set of apparently euchromatic *Drosophila* genes were identified that may be targets for dose-dependent HP1 regulation and can be affected by known modifiers of *Drosophila* PEV (Hwang et al. 2001). However, it is unclear if they variegate and/or what the nature of their chromatin structure is. A common theme among these and other studies of PEV is that variegation is extremely sensitive to HP1 dosage (Eissenberg et al. 1992; Hwang et al. 2001). It will be important to determine if this HP1 dose sensitivity holds true for repression of a mammalian euchromatic gene and if it can influence the stable component demonstrated in our system.

The interplay between histone methylation and DNA methylation

Our comparative ChIP analysis of silent and active clones clearly shows enrichment of KAP1, SETDB1, HP1, and H3-MeK9 at the silenced luciferase transgene promoter, all of which are likely to contribute to the stably silenced phenotype. However, most remarkable is the finding of increased DNA methylation in the stably silenced clone also centered on the proximal promoter region bound by these proteins. That this DNA methylation contributes to the silencing is suggested by the potent reactivation of the silent locus by the combined action of 5AC + TSA. This observation suggests that H3-MeK9 methylation and HP1 recruitment on a euchromatic gene may ultimately lead to DNA methylation. Identifying which component of the histone-directed silencing machinery serves as the signal for recruitment and/or activation of the DNA methyltransferases required for carrying out the DNA modification will be critical for understanding this system. Moreover, once the DNA methylation mark is established, it likely plays an active role in maintaining the histone-directed machinery at the silent locus. This could be accomplished in two ways. First, the MBD2/3 component of the NuRD histone deacetylase complex could directly bind the methylated DNA and maintain HDAC activity at the locus. Second, in addition to the catalytic SET domain, the SETDB1 protein also encodes a CpG DNA methyl binding domain, which, if functional, would maintain the H3-MeK9 activity at the locus (Schultz et al. 2002). That this may occur is supported by our preliminary observation that 5AC + TSA reactivation is accompanied by rapid loss of SETDB1 and H3-MeK9 at the locus as assessed by ChIP assays (K. Ayyanathan and F.J. Rauscher, unpubl.). In summary, the development of this tissue culture model will provide a valuable system to study in detail the sequence of events and the coordinated interplay between histone methylation, HP1 deposition, and DNA methylation.

Materials and methods

Expression and reporter plasmids

The construction of the pcKRAB–PAX3 (pcKP) plasmid has been described (Ayyanathan et al. 2000). The pcKRAB–PAX3–HBD (KPHBD) plasmid was constructed by fusing the estrogen receptor (ERHBD^{TM}) in frame with the COOH terminus of the PAX3 DNA binding domain (DBD) at a unique *Eco*RI site. The pcKRAB(DV)–PAX3–HBD (K(DV)PHBD) plasmid was constructed by subcloning a mutant KRAB ($DV_{18,19}AA$) domain (Margolin et al. 1994) as a *Hin*dIII/*Bam*HI fragment into pcKRAB–PAX3–HBD plasmid, thus replacing the wild-type KRAB domain. A similar strategy was utilized to construct the pcSNAG–PAX3–HBD (SPHBD), pcPLZF(POZ)–PAX3–HBD (PPHBD), pcEngrailed–PAX3–HBD (EPHBD), and pcWT1–PAX3–HBD (WPHBD) expression plasmids. The pcVP16–PAX3 (VPDBD) expression plasmid was constructed by inserting the VP16 acidic activation domain (residues 1–98) using a similar approach. All PCR-derived nucleotide sequences and the appropriate fusion junctions were confirmed by sequencing both DNA strands. A *Pvu*II fragment (zeocin^R cassette) from pcDNA3.1/Zeo plasmid (Invitrogen) was cloned into a unique *Pvu*II site in the CD19–TK–LUC plasmid to generate the CD19–TK–LUC–Zeo^R plasmid (Ayyanathan et al. 2000).

Antibodies

The sources of antibodies are: α-PAX3 IgG (Fredericks et al. 2000), affinity-purified α-KAP1 (raised against amino acids 20–418; Schultz et al. 2001), α-SETDB1 (Schultz et al. 2002), α-H3-MeK9 (Abcam, Inc.; this antibody recognizes di-methyl Lys 9), α-NT2 (raised against amino acids 323–345; Tanaka et al. 2002), and α-HP1α and α-HP1γ IgGs. The α-HP1α and α-HP1γ are mouse monoclonal antibodies, which were produced using 6-HIS-tagged, full-length human antigens. These reagents do not cross-react with the other human or mouse HP1 orthologs and will be described elsewhere (D.C. Schultz, M.S. Lechner, K. Ayyanathan, and F.J. Rauscher III, unpubl.).

Transient transfections and luciferase reporter assays

Protein expression was confirmed by transient transfection of COS-1 cells followed by immunoprecipitation of the [^{35}S]-L-methionine-labeled extracts with α-PAX3 IgG (Ryan et al. 1999; Ayyanathan et al. 2000). 4-OHT-dependent repression potentials were monitored via transient transfection as previously described (Ayyanathan et al. 2000). Cells were treated with vehicle (0.1% ethanol) or 500 nM 4-OHT (Research Biochemicals International). Fold repression was determined as the ratio of normalized light units in vehicle-treated cells versus 4-OHT treated cells.

For 4-OHT washout experiments, cells were seeded at 0.5×10^4 in 60-mm dishes. Duplicate dishes of cells were then

treated with either 500 nM 4-OHT (+OHT) or 0.1% ethanol (–OHT) for the indicated time. At the end of induction, one dish of –OHT and +OHT-treated cells were harvested while the remaining dishes were subjected to three washes each day. On fourth day, the cells were harvested by trypsinization and replated into new dishes. Washings were continued for an additional 4 d. Cells were harvested daily and normalized luciferase activities determined.

Generation and characterization of stably transfected cell clones

NIH3T3 cell lines containing a stably expressed chimeric repressor and the CD19–TK–LUC–Zeo^R reporter plasmid were generated by cotransfection and selection in growth medium containing 500 µg/mL G418 + 100 µg/mL zeocin. Colonies were ring-cloned, expanded, and tested for luciferase activity. The luciferase activities were normalized to protein concentration and expressed as light units/O.D. at A_{595}. Stable expression of the chimeric RD–PAX3–HBD proteins in clones was tested by immunoprecipitation with α-PAX3.

Quantitative RT–PCR (Q-RT–PCR)

Total RNAs were isolated using TRIzol Reagent (Life Technologies) and oligo-dT primed first-strand cDNAs produced. Luciferase, neomycin^R, zeocin^R, NT2, Col11a2, and RXRβ mRNAs were amplified by PCR for the number of cycles indicated. The respective primer pairs were: **LUC2** (5′-CAAGGATAT GGGCTCAC-3′) and **LUC3** (5′-GACCTTTCGGTACTTCG-3′), **NEO1** (5′-TCAGCGCAGGGGCGCCCGGTTCTTT-3′) and **NEO2** (5′-ATCGACAAGACCGGCTTCCATCCGA-3′), **ZEO1** (5′-ATGGCCAAGTTGACCAG-3′) and **ZEO2** (5′-TCAGTCC TGCTCCTCG-3′), **NT2-1** (5′-GCCAGGCTAGAAGGGAGG-3′) and **NT2-2** (5′-GGTGTCTGTTGAGGTTGG-3′), **COL1** (5′-GGCCTCAGCCTAGCAGATGG-3′) and **COL2** (5′-GGCTTA TGAAGTCTTGCTGG-3′), **RXR1** (5′-GGCTCTGTGCAATC TGCGGG-3′) and **RXR2** (5′-GTCCACAGGCATCTCCTCAG GG-3′). Following electrophoresis and Southern blotting, the images were quantified using ImageQuant (Molecular Dynamics).

In vivo analysis of chromatin structure

Nuclei were prepared essentially as described (Mymryk et al. 1997). For the restriction enzyme accessibility assays, the nuclei were diluted into appropriate 1× restriction enzyme buffers in a 500-µL reaction volume and digested with 250 units of the indicated enzymes for 10 min at room temperature. The reaction was terminated by adding 100 µg of proteinase-K in 10 mM Tris-HCl at pH 8.0, 10 mM EDTA, 10 mM NaCl, and incubated at 37°C for 14–16 h. These samples were phenol:chloroform extracted and ethanol precipitated. The purified DNA was digested with *Bam*HI to completion and purified. Equal amounts of the DNA samples were taken for reiterative, primer extension PCR reactions with the following ^{32}P-labeled primers: **LUC1** (5′-TCCAGGAACCAGGGCGTATCTCT-3′) or **ZEO2** or **NEO2**. Primer extension products were extracted with phenol-:chloroform and ethanol precipitated. Dried DNA pellets were dissolved in formamide gel-loading buffer and electrophoresed in 7M Urea/5% acrylamide gels in TBE buffer. The gels were fixed in 10% acetic acid, dried, and autoradiographed.

ChIP

Soluble, sonicated chromatin was prepared essentially as described (Schultz et al. 2002). Chromatin fractions were immunoprecipitated with 10 µg of indicated antibodies and the immune complexes were recovered using protein-A sepharose beads and processed as described (Schultz et al. 2002). The immunoprecipitated and input DNA samples were used in quantitative PCR using the following oligonucleotides: **LUC1**, **LUC2**, **LUC3**, **ZEO2**, **PBS3** (5′-GAATACACGGAATTGGA TCCG-3′), **PBS1** (5′-GATCGATAATTCGAGCTACTG-3′), **PBS2** (5′-GAGCTCGGTACCCGGGTCG-3′), **TKP1** (5′-GCGC GGTCCCAGGTCCACTT-3′), **SVP1** (5′-CCAGTTCCGCCCA TTCTCCC-3′), **C1** (5′-GGATGCTGCCACGGCCTGAGG-3′), **C2** (5′-GGGTCTGCCAGGAGCCTGTGG-3′), **C3** (5′-GGGTC GCTATCTATAGCTGG-3′), **C4** (5′-GTCCTTTCACACCAC GGCAG-3′), **C5** (5′-GGCCTCAGCCTAGCAGATGG-3′), **C6** (5′-GAGCACTCAGACCTTCCAGAGG-3′), **C7** (5′-GGAACATCA GGTGGTGACGG-3′), **C8** (5′-GCAGCCCATCCTTCCCTGC AGG-3′). The PCR products were resolved on 1.5% agarose gels and the bands were quantified using IQMac v1.2 software.

FISH

For detection of the luciferase gene, either 0.1% ethanol (-OHT) or 500 nM 4-OHT (+OHT) treated KPHBD21 cells were fixed in 4% paraformaldehyde, permeabilized in 0.2% Triton X-100, and then treated with RNase. After equilibration in 2× SSC, cells were dehydrated in an ethanol series (70%, 80%, and 100%). The CD19–TK–LUC plasmid was labeled with biotin-16-dUTP by nick-translation and dissolved at 10 ng/µL in 50% formamide in 2× SSC containing 10% dextran sulfate, 100 ng/µL salmon sperm DNA, 1 µg/µL yeast tRNA, and 1.5 µg/µL Cot-1 DNA. Both the probe and the cells were simultaneously heated at 91°C for 4 min to denature DNA and incubated overnight at 37°C. After hybridization, specimens were serially washed at 37°C with 50% formamide in 2× SSC, 2× SSC, and 0.25× SSC. Hybridized probes were detected with FITC-streptavidin (Vector Laboratories). Finally, cells were equilibrated in PBS, stained for DNA with either DAPI or Hoechst (2 ng/mL) and mounted in Fluoromount G (Southern Biotechnology Associates, Inc). Cells were analyzed with a Leitz Fluovert inverted microscope equipped with a digital camera. Images were obtained using QED Imaging software.

Clonal analysis of luciferase gene expression

Three independent subclones (KPHBD21-8, KPHBD21-39, and KPHBD21-49) that express high-level luciferase activities were treated with either 500 nM 4-OHT (+OHT) or 0.1% ethanol (–OHT) for 4 d continuously and subjected to limiting dilution cloning in 96-well plates. Wells containing single cells were identified and the cells were propagated for ~40 doublings in normal growth medium. Whole-cell lysates were prepared from these individual progeny and normalized luciferase activities were determined.

DNA methylation analysis

MSPCR: Genomic DNAs were extracted from the silent (cl 39–40) and active (cl 39–45) clones, and 1 µg of each was treated with sodium bisulfite to convert the unmethylated cytosines to uracil as described (Herman et al. 1996). These DNAs were PCR amplified using unmethylated sense (US; 5′-GTTGATTTGGG TATTGAGTTTGAG-3′) or methylated sense (MS; 5′-GTCGA TTCGGGTATCGAGTTCG-3′) primers (present in the TK promoter) and an antisense primer (UMA1; 5′-GGTTTTATTTTT TAGAGGATAGAATGG-3′; present near the start of luciferase cDNA). The PCR products were electrophoresed on a 1.5% agarose gel and photographed.

Sodium bisulfite-genomic sequencing: The sodium bisulfite-modified DNAs were PCR amplified using UMS1 (sense; 5′-GTTTTAGTGTTTTATGTTTTAGG-3′) and UMA1 (antisense) primers that were present in regions free of any CpG residues (to avoid any preferential amplification). The PCR products were TA-cloned into the pCR II vector. At least 18 recombinant clones of each were sequenced in both directions. Results are presented for seven representative clones in each category.

Acknowledgments

We thank William Fredericks for proofreading the manuscript. We thank Trevor Littlewood and Gerard Evan for the ERHBD™ plasmid. We thank Nick Barlev and Shelley Berger for the ChIP protocol. We thank Jim Herman for advice on MSPCR and Bisulfite-genomic sequencing. We thank Andrey Loboda and Michael Showe for statistical analysis. K.A. is supported by CA095620. D.C.S. is supported by DAMD17-98-1-8269. M.S.L. is supported by CA09171. G.G.M. is supported by AI41136, AI52742, AI53237, and GM64691. F.J.R. is supported in part by National Institutes of Health grants, Core grant CA10815, CA92088, CA095561, DAMD17-96-1-6141, DAMD17-02-1-0631, the Irving A. Hansen Memorial Foundation, the Susan G. Komen Breast Cancer Foundation, and the Emerald Foundation. Support from Commonwealth Universal Research Enhancement Program, Pennsylvania Department of Health, is gratefully acknowledged.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Abrink, M., Ortiz, J.A., Mark, C., Sanchez, C., Looman, C., Hellman, L., Chambon, P., and Losson, R. 2001. Conserved interaction between distinct Kruppel-associated box domains and the transcriptional intermediary factor 1 beta. *Proc. Natl. Acad. Sci.* **98:** 1422–1426.

Ayyanathan, K., Fredericks, W.J., Berking, C., Herlyn, M., Balakrishnan, C., Gunther, E., and Rauscher III, F.J. 2000. Hormone-dependent tumor regression in vivo by an inducible transcriptional repressor directed at the PAX3–FKHR oncogene. *Cancer Res.* **60:** 5803–5814.

Baker, W.K. 1968. Position-effect variegation. *Adv. Genet.* **14:** 133–169.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410:** 120–124.

Cameron, E.E., Bachman, K.E., Myohanen, S., Herman, J.G., and Baylin, S.B. 1999. Synergy of demethylation and histone deacetylase inhibition in the re- expression of genes silenced in cancer. *Nat. Genet.* **21:** 103–107.

Eissenberg, J.C., Morris, G.D., Reuter, G., and Hartnett, T. 1992. The heterochromatin-associated protein HP-1 is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation. *Genetics* **131:** 345–352.

Festenstein, R., Sharghi-Namini, S., Fox, M., Roderick, K., Tolaini, M., Norton, T., Saveliev, A., Kioussis, D., and Singh, P. 1999. Heterochromatin protein 1 modifies mammalian PEV in a dose- and chromosomal-context-dependent manner. *Nat. Genet.* **23:** 457–461.

Fredericks, W.J., Ayyanathan, K., Herlyn, M., Friedman, J.R., and Rauscher III, F.J. 2000. An engineered PAX3–KRAB transcriptional repressor inhibits the malignant phenotype of alveolar rhabdomyosarcoma cells harboring the endogenous PAX3–FKHR oncogene. *Mol. Cell. Biol.* **20:** 5019–5031.

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., and Rauscher III, F.J. 1996. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes & Dev.* **10:** 2067–2078.

Gray, S. and Levine, M. 1996. Short-range transcriptional repressors mediate both quenching and direct repression within complex loci in *Drosophila*. *Genes & Dev.* **10:** 700–710.

Herman, J.G., Graff, J.R., Myohanen, S., Nelkin, B.D., and Baylin, S.B. 1996. Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. *Proc. Natl. Acad. Sci.* **93:** 9821–9826.

Hwang, K.K., Eissenberg, J.C., and Worman, H.J. 2001. Transcriptional repression of euchromatic genes by *Drosophila* heterochromatin protein 1 and histone modifiers. *Proc. Natl. Acad. Sci.* **98:** 11423–11427.

James, T.C. and Elgin, S.C. 1986. Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. *Mol. Cell. Biol.* **6:** 3862–3872.

Jaynes, J.B. and O'Farrell, P.H. 1991. Active repression of transcription by the engrailed homeodomain protein. *EMBO J.* **10:** 1427–1433.

Jenuwein, T. and Allis, C.D. 2001. Translating the histone code. *Science* **293:** 1074–1080.

Kirschmann, D.A., Lininger, R.A., Gardner, L.M., Seftor, E.A., Odero, V.A., Ainsztein, A.M., Earnshaw, W.C., Wallrath, L.L., and Hendrix, M.J. 2000. Down-regulation of HP1Hsα expression is associated with the metastatic phenotype in breast cancer. *Cancer Res.* **60:** 3359–3363.

Kuo, M.H. and Allis, C.D. 1998. Roles of histone acetyltransferases and deacetylases in gene regulation. *Bioessays* **20:** 615–626.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410:** 116–120.

Lechner, M.S., Begg, G.E., Speicher, D.W., and Rauscher III, F.J. 2000. Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: Direct chromoshadow domain-KAP-1 corepressor interaction is essential. *Mol. Cell. Biol.* **20:** 6449–6465.

Li, X., Lopez-Guisa, J.M., Ninan, N., Weiner, E.J., Rauscher III, F.J., and Marmorstein, R. 1997. Overexpression, purification, characterization, and crystallization of the BTB/POZ domain from the PLZF oncoprotein. *J. Biol. Chem.* **272:** 27324–27329.

Littlewood, T.D., Hancock, D.C., Danielian, P.S., Parker, M.G., and Evan, G.I. 1995. A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. *Nucleic Acids Res.* **23:** 1686–1690.

Locke, J., Kotarski, M.A., and Tartof, K.D. 1988. Dosage-dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* **120:** 181–198.

Lu, B.Y., Ma, J., and Eissenberg, J.C. 1998. Developmental regulation of heterochromatin-mediated gene silencing in *Drosophila*. *Development* **125:** 2223–2234.

Madden, S.L., Cook, D.M., Morris, J.F., Gashler, A., Sukhatme, V.P., and Rauscher III, F.J. 1991. Transcriptional repression mediated by the WT1 Wilms tumor gene product. *Science* **253:** 1550–1553.

Margolin, J.F., Friedman, J.R., Meyer, W.K., Vissing, H., Thiesen, H.J., and Rauscher III, F.J. 1994. Kruppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci.* **91:** 4509–4513.

Minc, E., Courvalin, J.C., and Buendia, B. 2000. HP1gamma associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes. *Cytogenet. Cell. Genet.* **90:** 279–284.

Morrison, A.J., Sardet, C., and Herrera, R.E. 2002. Retinoblastoma protein transcriptional repression through histone deacetylation of a single nucleosome. *Mol. Cell. Biol.* **22:** 856–865.

Mymryk, J.S., Fryer, C.J., Jung, L.A., and Archer, T.K. 1997. Analysis of chromatin structure in vivo. *Methods* **12:** 105–114.

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292:** 110–113.

Nielsen, A.L., Oulad-Abdelghani, M., Ortiz, J.A., Remboutsika, E., Chambon, P., and Losson, R. 2001. Heterochromatin formation in mammalian cells: Interaction between histones and HP1 proteins. *Mol. Cell* **7:** 729–739.

Pelengaris, S., Littlewood, T., Khan, M., Elia, G., and Evan, G. 1999. Reversible activation of c-Myc in skin: Induction of a complex neoplastic phenotype by a single oncogenic lesion. *Mol. Cell* **3:** 565–577.

Peng, H., Begg, G.E., Schultz, D.C., Friedman, J.R., Jensen, D.E., Speicher, D.W., and Rauscher III, F.J. 2000. Reconstitution of the KRAB–KAP-1 repressor complex: A model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein–protein interactions. *J. Mol. Biol.* **295:** 1139–1162.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406:** 593–599.

Ryan, R.F., Schultz, D.C., Ayyanathan, K., Singh, P.B., Friedman, J.R., Fredericks, W.J., and Rauscher III, F.J. 1999. KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: A potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol. Cell. Biol.* **19:** 4366–4378.

Schultz, D.C., Friedman, J.R., and Rauscher III, F.J. 2001. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: The PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. *Genes & Dev.* **15:** 428–443.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher III, F.J. 2002. SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes & Dev.* **16:** 919–932.

Seum, C., Delattre, M., Spierer, A., and Spierer, P. 2001. Ectopic HP1 promotes chromosome loops and variegated silencing in *Drosophila*. *EMBO J.* **20:** 812–818.

Smothers, J.F. and Henikoff, S. 2001. The hinge and chromo shadow domain impart distinct targeting of HP1-like proteins. *Mol. Cell. Biol.* **21:** 2555–2569.

Tanaka, K., Tsumaki, N., Kozak, C.A., Matsumoto, Y., Nakatani, F., Iwamoto, Y., and Yamada, Y. 2002. A Kruppel-associated box-zinc finger protein, NT2, represses cell-type-specific promoter activity of the α 2(XI) collagen gene. *Mol. Cell. Biol.* **22:** 4256–4267.

Wakimoto, B.T. 1998. Beyond the nucleosome: Epigenetic aspects of position-effect variegation in *Drosophila*. *Cell* **93:** 321–324.

Wallrath, L.L. 1998. Unfolding the mysteries of heterochromatin. *Curr. Opin. Genet. Dev.* **8:** 147–153.

Wolffe, A.P. and Hansen, J.C. 2001. Nuclear visions: functional flexibility from structural instability. *Cell* **104:** 631–634.

Wu, J. and Grunstein, M. 2000. 25 years after the nucleosome model: Chromatin modifications. *Trends Biochem Sci.* **25:** 619–623.

Zhang, Y. and Reinberg, D. 2001. Transcription regulation by histone methylation: Interplay between different covalent modifications of the core histone tails. *Genes & Dev.* **15:** 2343–2360.

Zweidler-Mckay, P.A., Grimes, H.L., Flubacher, M.M., and Tsichlis, P.N. 1996. Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. *Mol. Cell. Biol.* **16:** 4024–4034.

 Related Commentary, page 304

A proximal activator of transcription in epithelial-mesenchymal transition

Christo D. Venkov,[1] Andrew J. Link,[2] Jennifer L. Jennings,[2] David Plieth,[1] Tsutomu Inoue,[1] Kojiro Nagai,[1] Carol Xu,[1] Yoana N. Dimitrova,[3] Frank J. Rauscher III,[4] and Eric G. Neilson[1,5]

[1]Department of Medicine, [2]Department of Microbiology and Immunology, and [3]Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. [4]Wistar Institute, Philadelphia, Pennsylvania, USA.[5]Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, Tennessee, USA.

Epithelial-mesenchymal transition (EMT) is an important mechanism for phenotypic conversion in normal development and disease states such as tissue fibrosis and metastasis. While this conversion of epithelia is under tight transcriptional control, few of the key transcriptional proteins are known. Fibroblasts produced by EMT express a gene encoding fibroblast-specific protein 1 (FSP1), which is regulated by a proximal *cis*-acting promoter element called fibroblast transcription site–1 (FTS-1). In mass spectrometry, chromatin immunoprecipitation, and siRNA studies, we used FTS-1 as a unique probe for mediators of EMT and identified a complex of 2 proteins, CArG box–binding factor–A (CBF-A) and KRAB-associated protein 1 (KAP-1), that bind this site. Epithelial cells engineered to conditionally express recombinant *CBF-A* (*rCBF-A*) activate the transcription of *FSP1* and undergo EMT. The FTS-1 response element also exists in the promoters modulating a broader EMT transcriptome, including *Twist*, and *Snail*, as well as *E-cadherin*, β-*catenin*, *ZO 1*, *vimentin*, α*1(I) collagen*, and α–*smooth muscle actin*, and the induction of *rCBF-A* appropriately alters their expression as well. We believe formation of the CBF-A/KAP-1/FTS-1 complex is sufficient for the induction of *FSP1* and a novel proximal activator of EMT.

Introduction

The mechanisms governing molecular signals for epithelial-mesenchymal transition (EMT) are increasingly more complex (1, 2). Fibrogenesis during wound healing or following organ inflammation depends on the formation and proliferation of new fibroblasts by EMT. We previously described a gene encoding fibroblast-specific protein 1 (FSP1) that activates in epithelia during EMT and is constitutively and selectively present in newly formed fibroblasts thereafter; FSP1, also known as S100A4 in the cancer literature, is an intracellular calcium-binding protein whose appearance is linked to EMT (3–7), tissue fibrosis (4, 8), pulmonary vascular disease (9), increased tumor cell motility and invasiveness (10), and metastatic tumor development (11–16). FSP1 helps epithelia transition to new morphology and motility based on its ability to influence levels of intracellular calcium and actin disassembly when transfected into cultured cells. The important role of FSP1 in EMT is underscored by findings that induction of EMT in vitro by epithelial growth factor (EGF) and TGF-β is blocked by antisense oligomers against mRNA encoding *FSP1* (6) and that levels of mRNA encoding *E-cadherin* are inversely correlated with *FSP1* expression in invasive lines of squamous cell carcinoma (17).

The transcriptional control of EMT diversifies the lineage specification of epithelia during development (18), lineage commitment of epithelia in tissue fibrosis (1), and the stimulation of tumor metastasis (2, 11, 19). An increasing number of transcription factors appear to activate EMT in various settings, including Snail (20), Twist (11), high mobility group A2 (HMGA2) (21), Slug (16), and Ets-1 (22, 23). However, the set of transcription factors controlling EMT is far from complete.

In an earlier report using promoter deletion constructs for *FSP1*, we identified a novel *cis*-acting element in the *FSP1* promoter called fibroblast transcription site–1 (FTS–1), which activates transfected reporters in fibroblasts and forms specific complexes in EMSAs with nuclear extracts (24). Using both EMSA and DNA affinity chromatography followed by mass spectrometry analysis (25), we identified 2 key transcriptional proteins in this complex, namely the CArG box–binding factor–A (CBF-A) (26) and the KRAB-associated protein 1 (KAP-1) (27). These proteins form a ternary complex with FTS-1 to engage the transcription of *FSP1* and play a key functional role in forming fibroblasts by EMT.

Results

Purification of the FTS-1–binding complex. In our original report of FTS-1 (24), we used a 100-bp oligonucleotide from the proximal *FSP1* promoter to identify the fibroblast-specific complex by EMSA. To minimize the background while ensuring specificity of binding to the 5-nucleotide FTS-1 motif, this time we used a 25-bp sequence containing the core pentanucleotide (TTGAT) with adjacent promoter sequences. Complexes identified by EMSA using both oligonucleotides and fibroblast nuclear extracts had a similar number and pattern of distribution when run in parallel (Figure 1A); differences in the relative mobility of the bands arose from the different lengths of the oligonucleotides. The 4 complexes were competed to varying degree by challenge with 20× molar excess unlabeled probe. The specificity of the complexes formed with the 25-bp FTS-1 oligonucleotide was further tested by competition analysis

Nonstandard abbreviations used: CBF-A, CArG box–binding factor–A; ChIP, chromatin immunoprecipitation; Ct, threshold cycle; EEF1A1, eukaryotic elongation factor 1 α1; EGF, epithelial growth factor; EMT, epithelial-mesenchymal transition; FSP1, fibroblast-specific protein 1; FTS-1, fibroblast transcription site–1; HMGA2, high mobility group A2; KAP-1, KRAB-associated protein 1; LEF-1, lymphoid enhancing binding factor; mIMCD, mouse inner medullary collecting duct; qRT-PCR, quantitative real-time PCR; *rCBF-A*, recombinant *CBF-A*; SRE, serum response factor; UUO, unilateral ureteral obstruction; YB-1, Y-box transcription factor.

Conflict of interest: The authors have declared that no conflict of interest exists.

Citation for this article: *J. Clin. Invest.* **117**:482–491 (2007). doi:10.1172/JCI29544.

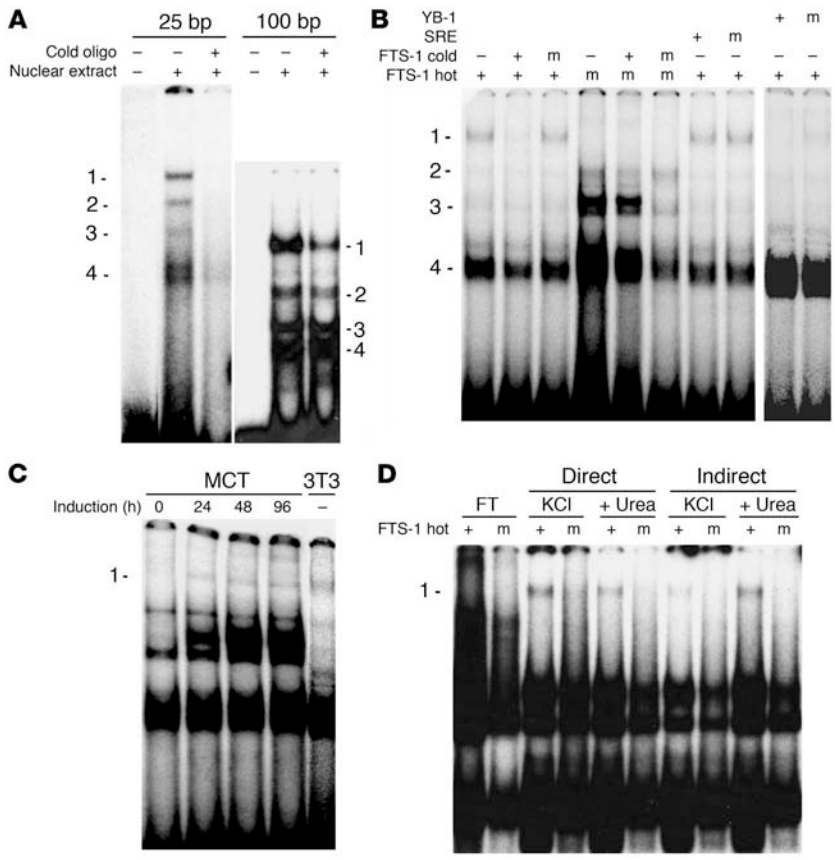


Figure 1
Identification and purification of the FTS-1 complex by EMSA and DNA affinity chromatography. (**A**) Similar complexes form in EMSA with fibroblast nuclear protein and either 25-bp or 100-bp probes, both containing the FTS-1 core pentanucleotide. The complexes were challenged with nonlabeled probes. (**B**) Specificity of the FTS-1 complex established by EMSA following challenge by various nonlabeled competitors. Fibroblast nuclear protein was incubated with the 25-bp radiolabeled FTS-1 or its mutant form (m). The complexes were challenged with DNA probes specific for different regulatory sites. The complex specific to the TTGAT pentanucleotide (marked 1) is at the top; (m) indicates the mutated oligonucleotides. (**C**) The FTS-1 complex (marked 1) is formed by MCT cells induced to EMT as evidenced by EMSA at different times after induction. A fibroblast-derived FTS-1 complex is shown as a positive control. (**D**) DNA affinity-purified proteins form only 1 complex with FTS-1 (marked 1) in EMSA. Fractions purified by the direct or indirect approach were eluted with 0.5 M KCl followed by 0.5 M KCl plus 5 M urea, dialyzed against 0.1 M KCl, and incubated with radiolabeled FTS-1 or its mutated form (m). The flow-through (FT) from the direct incubation could not form the FTS-1 complex.

using related and unrelated probes (Figure 1B). Complexes formed with radiolabeled FTS-1 or its mutant (CCAGC instead of TTGAT) were competed at 20× molar excess with unlabeled species or with consensus sites for the Y-box transcription factor (YB-1) and serum response factor (SRE). The YB-1 core sequence differs in 1 nucleotide from FTS-1 (TTGGT versus TTGAT), and SRE is unrelated to FTS-1 and contains a CarG box DNA motif. Mutant constructs of these oligonucleotides were also used in the competition analysis. Of the 4 complexes formed by the FTS-1 probe, the more slowly migrating complex 1 was not competed by its mutant and the only one not formed by a radiolabeled mutant. Complex 1 was competed by YB-1 because of structural similarity, but when used as a probe, YB-1 did not form a similar complex in EMSA (data not shown). Complex 1 was also unaffected by challenge with the YB-1 mutant or SRE. Competition analysis confirmed this complex as specific for the core FTS-1 site (TTGAT). Unlabeled competitors did not affect complex 4. Complexes 2 and 3, although not competed by YB-1 or by SRE, were formed with both FTS-1 and its mutant (Figure 1B). Therefore, only complex 1 was specific for FTS-1. Complex 1 also formed in EMSA when the probe was incubated with nuclear extracts from kidney epithelial cells undergoing EMT (Figure 1C) in parallel with the expression of FSP1 (6).

These results were confirmed in another epithelial cell line from mouse inner medullary collecting duct (mIMCD; ref. 28). mIMCD cells stimulated to undergo EMT with TGF-β/EGF formed FTS-1 complexes on day 5 (data not shown).

The FTS-1–binding complex identified by EMSA was purified independently using DNA affinity chromatography. We used a biotinylated tetramer of the 25-bp FTS-1 sequences attached to streptavidin-coated magnetic beads. The binding reactions were performed under the same conditions as the EMSAs to ensure comparability. Initially we used 2 experimental variants: a direct incubation of nuclear extracts with the FTS-1 oligonucleotide; and an indirect approach using the flow-through from an initial incubation with the mutant tetramer for a secondary incubation with the TTGAT tetramer. Proteins that bind specifically were eluted with high salt, with or without urea, and analyzed by EMSA using radiolabeled mutant or wild-type FTS-1 (Figure 1D). The results of this analysis demonstrate that affinity-purified proteins were able to form a DNA-protein complex, which comigrated with complex 1. There was no difference in the complexes formed by the proteins purified by either approach, and subsequently we used only the indirect approach to minimize contaminating proteins. These results established that both EMSA and DNA affinity chromatography consistently identified the same FTS-1 complex. The complex recovered both from EMSA and from DNA affinity purification was used separately for protein identification utilizing tandem mass spectrometry (25).

Proteomic analysis of FTS-1 complexes. The combination of EMSA with mass spectrometry identified several protein signatures. Two independent peptides from the FTS-1 complex (Table 1) were identified as CBF-A (26). Two other peptides belonged to the tripartite motif protein 28 (TIF1β/Trim28), a mouse homolog of KAP-1 (27). Other identified peptides were from the eukaryotic elongation factor 1 α1 (EEF1A1) and the DNA replication-licensing factor MCM3.

A second technique, DNA affinity chromatography followed by mass spectrometry, also identified CBF-A and KAP-1 in the FTS-1 complex. Three independent CBF-A peptides were identified;

Table 1
Peptide identification in the FTS-1–binding complex

Protein	Accession no.	Peptide sequence	
		Purification by EMSA	Purification by DNA affinity chromatography
CBF-A	Q99020	GFVFITFKEEDPVK, IFVGGLNPEATEEK	GFVFITFKEEDPVK, IFVGGLNPEATEEK, EVYQQQQYGSGGR
Tripartite motif protein 28 (TIF1β/Trim28)/KAP-1	Q62318	EEDGSLSLDGADSTGVVAK, LDLDLTSDSQPPVFK	EEDGSLSLDGADSTGVVAK, LDLDLTSDSQPPVFK
EEF1A1 (EF-1α1/EF-Tu)	P10126	EHALLAYTLGVK, IGGIGTVPVGR, RYEEIVK, STTTGHLIYK, THINIVVIGHVDSGK, YYVTIIDAPGHR	None

2 are shared by both purification approaches, and a third one, EVYQQQQYGSGGR, appeared in the affinity-purified complex (Table 1). The 2 KAP-1 peptides described above were also confirmed. However, neither EEF1A1 nor MCM3 peptides were present in the affinity-purified complex.

We used 2-dimensional EMSA/SDS-PAGE followed by immunoblotting to further characterize the composition of the FTS-1 complex. To this end, we ran nuclear extracts in EMSA with and without incubation with an FTS-1 probe and then subjected the gel pieces to second-dimension SDS-PAGE. After subsequent transfer to membrane and immunostaining, we confirmed the presence of CBF-A and KAP-1 in the FTS-1 complex (Figure 2A). Immunostaining indicated the absence of CBF-A and KAP-1 in control extracts, as shown on the left side of each panel. Only MCM3-positive immunostaining was detected on the left side of the panel (data not shown), which suggests this latter protein does not form a complex with FTS-1 and only comigrates with the complex in the first dimension. An antibody against EEF1A1 did not detected it either in the complex or as a comigrating protein, suggesting it was an artifact.

The binding of CBF-A and KAP-1 to the FTS-1 probe was confirmed additionally by EMSA supershifts (Figure 2B). The CBF-A antibody supershifts the FTS-1 complex, while antibody against KAP-1 abolishes it. We suspect the KAP-1 antibody most likely inhibits sterically the interaction of CBF-A/KAP-1 with FTS-1. Anti-rabbit IgG did not affect the mobility or integrity of the FTS-1 complex. Finally, we performed a chromatin immunoprecipitation (ChIP) assay with the CBF-A antibody using a rabbit preimmune serum as a negative control. Western blots of precipitates (Figure 2C) confirmed the presence of both CBF-A and KAP-1 as participating in a complex with FTS-1. The increased expression of CBF-A and KAP-1 in cell nuclei during EMT was further demonstrated by immunostaining MCT epithelia stimulated with TGF-β and EGF in culture (6); the nuclear staining for CBF-A increased from 14.8% of cells in unstimulated epithelia to 47.9% in EMT-induced fibroblasts, and nuclear KAP-1 levels increased from 3.3% to 30.2% (Figure 2D). These data collectively suggest that CBF-A and KAP-1 are components that recognize the FTS-1 element in cell nuclei.

rCBF-A induces de novo expression of FSP1 in kidney epithelial cells. Differentiated epithelial cells do not express FSP1 nor form nuclear complexes with FTS-1 (24). When they undergo EMT, however, they form FTS-1 complexes (Figure 1C) and synthesize FSP1 (6). To understand the functional significance of CBF-A in the induction of *FSP1,* we expressed recombinant CBF-A (rCBF-A) in epithelial cells. For this purpose, the kidney epithelial cell line MCT was stably transfected with a TET-inducible vector for conditional expression of *rCBF-A* (MCT/rCBF-A). We selected several clones that displayed controlled overexpression of *rCBF-A* following exposure to doxycycline. We refer to the noninduced state as MCT/rCBF-A$^-$ and to the doxycycline-induced state as MCT/rCBF-A$^+$. The induction of *rCBF-A* with doxycycline led to the expression of *FSP1* transcripts (Figure 3A) and the coordinate appearance of FSP1 protein (Figure 3D). Overexpression of *rCBF-A* resulted in the formation of the FTS-1 complex, as confirmed by ChIP assay with CBF-A and KAP-1 antibodies (Figure 3B). Noninduced cells maintain low levels of endogenous CBF-A, and the FTS-1 DNA site could not be detected by PCR. However, cells induced with doxycycline formed FTS-1 complexes robustly (Figure 3B). Chromatin from 3T3 fibroblasts was immunoprecipitated in parallel as a positive control (Figure 3B, left), with comparable results. Furthermore, nuclear extracts from MCT/rCBF-A$^+$ cells, but not from wild-type MCT cells, formed the FTS-1 complex in EMSA with affinity similar to that of extracts from 3T3 fibroblasts, as demonstrated by challenge with increasing molar amounts of unlabeled competitor (Figure 3C). These results indicate that overexpression of *rCBF-A* leads to formation of the FTS-1 complex and activation of *FSP1* expression in epithelia in much the same way as *FSP1* is expressed by fibroblasts.

We further confirmed these findings in an in vivo kidney model of EMT and fibrosis following unilateral ureteral obstruction (UUO) (4). Obstructed kidneys undergoing EMT and fibrosis after 12 days (data not shown) had increased levels of CBF-A, FSP1, and α-SMA ($P < 0.002$; Figure 3E). This increase in CBF-A and FSP1 was accompanied by an expected decrease in E-cadherin expression. The expression of the other protein in the FTS-1 complex, KAP-1, was slightly upregulated in the UUO kidney, but the increase was not statistically significant.

Figure 3F demonstrates a change in the nuclear localization of CBF-A in fibrotic renal tissue 12 days after UUO in FSP1.GFP mice. CBF-A in the normal contralateral kidney was observed in the cytoplasm (speckled green) of tubular cells (nuclei counterstained red with propidium iodide); FSP1$^+$ fibroblasts expressing GFP under the control of the *FSP1* promoter were recolored dark blue in these images. In the frames labeled UUO1 and UUO2, the red nuclei have become yellow/orange with a change in localization of CBF-A from the cytoplasm (percentage of nuclei stained for CBF-A increased 21% in obstructed kidneys compared with 4% in contralateral controls; $P < 0.0001$), and the fibroblasts now stained light blue with the addition of CBF-A. This series of images suggests that the nuclei of many tubular cells in the UUO kidney are enriched for CBF-A and potentially positioned for an EMT event. The upregulation of CBF-A in UUO kidney established by quanti-

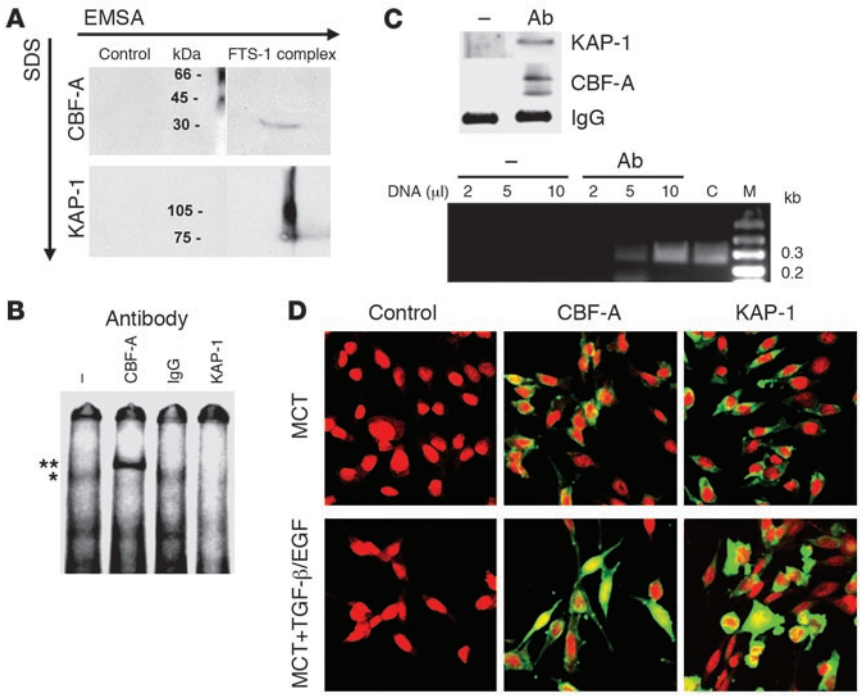


Figure 2

CBF-A and KAP-1 are integral parts of the FTS-1 complex. (**A**) Immunoblotting and 2D electrophoresis. First dimension: EMSA; second dimension: SDS-PAGE. Left: control for EMSA with nuclear protein without the probe; right (FTS-1 complex): protein with the FTS-1 probe. The absence of immunostaining on the left indicates that the investigated proteins do not comigrate with the complexes in the absence of DNA. (**B**) The participation of CBF-A and KAP-1 in the FTS-1 complex was confirmed by EMSA with their antibodies. The complex (*) was supershifted by the CBF-A antibody as indicated (**) and was abolished by anti–KAP-1 antibody; rabbit IgG did not affect the complex. (**C**) ChIP of 3T3 chromatin with the anti–CBF-A antibody or rabbit preimmune serum (–). Western blotting (upper panel): the immunoblots were probed with antibodies against KAP-1 and CBF-A; the stained IgG is shown as a loading control. PCR (lower panel): Increasing amounts (2, 5, and 10 μl) of DNA eluted from the immunoprecipitates were used as template in PCR with primers encompassing the FTS-1 site in both directions, producing a fragment of approximately 280 bp. Genomic DNA from 3T3 fibroblasts was used in the PCR as a positive control. M, DNA size ladder in kilobases (kb). (**D**) After induction of EMT, MCT cells showed increased nuclear colocalization of KAP-1 and CBF-A. MCT epithelia were stimulated for 5 days with TGF-β/EGF in culture. Cells were stained with nuclear propidium iodide dye (red) followed by indirect immunofluorescence with primary antibodies against CBF-A or KAP-1 or with control FITC–anti-rabbit IgG alone. Original magnification, ×630.

tative real-time PCR (qRT-PCR) analysis (Figure 3E) suggests that most of the nuclear protein is newly synthesized.

Overexpression of CBF-A engages the fibroblast transcriptome by initiating EMT. Treatment of cultured epithelia with a combination of EGF and TGF-β initiates EMT (6), which is associated with the activation of fibroblast markers such as FSP1, vimentin, and α-SMA and a downregulation of E-cadherin and the tight junction–associated protein zona occludens 1 (ZO-1); the induction of *FSP1* precedes the emergence of the EMT phenotype and the capacity of epithelia for movement, indicating a key early role for FSP1. A similar induction of FSP1 in MCT/rCBF-A$^+$ cells (Figure 3D) suggests that *rCBF-A* is sufficient to initiate EMT. We further checked for levels of expression of a number of other epithelial and fibroblast cell markers in MCT/rCBF-A$^+$ cells, as well as their cell motility and morphology. Results from immunoblotting indicate that MCT/rCBF-A$^+$ cells have increased levels of the fibroblast markers FSP1, α-SMA, vimentin, and α1(I) collagen, together with the newly identified markers of EMT, N-cadherin, Snail, and Twist (Figure 3D). The latter 2 proteins are transcription factors associated with increased motility of transitional epithelial cells (11, 20). The levels of E-cadherin, β-catenin, and ZO-1 were attenuated correspondingly. These changes in protein levels are indicative of cell remodeling associated with EMT (1, 2).

The differential expression of epithelial and fibroblast markers induced by *rCBF-A* was confirmed by changes in levels of mRNA (data not shown) and by immunostaining of MCT/CBF-A$^+$ compared with control epithelial cells 5 days after induction (Figure 4A). The pattern of expression and distribution of ZO-1, β-catenin, α-SMA, and vimentin in MCT/CBF-A$^+$ cells clearly indicates a change of phenotype associated with FSP1 expression. Furthermore, 5 days following the induction of *rCBF-A*, MCT/rCBF-A$^+$ cells but not control cells assumed a spindle-like, fibroblast morphology typical of epithelial transition (Figure 4B).

The change in phenotype induced by *rCBF-A* was associated with an increased ability of MCT/rCBF-A$^+$ fibroblasts to migrate directionally following serum stimulation in a Boyden chamber assay (Figure 4C). Conversely, inhibition of CBF-A expression in 3T3 fibroblasts by siRNA caused a coordinated downregulation of CBF-A and FSP1 proteins (Figure 5). Transiently transfected siRNA1 best inhibited *CBF-A* and *FSP1* expression after 48 hours by qRT-PCR from 3 independent experiments; 37% and 49%, respectively ($P \leq 0.001$). These data collectively indicate that overexpression of *rCBF-A* in epithelial cells leads to formation of the FTS-1 complex, expression of FSP1, modulation of epithelial and fibroblast genes including *Snail* and *Twist*, and subsequent EMT.

Discussion

In this study, we have discovered what we believe to be a new set of transcriptional proteins mediating early events in EMT. As an experimental probe, we used a novel promoter element (FTS-1) found in the gene encoding FSP1 to identify a novel complex of transcriptional proteins, CBF-A and KAP-1, which engage the genes encoding the EMT proteome. We conclude the following: first, CBF-A and KAP-1 can be found in a complex with the FTS-1 element in fibroblasts and in epithelia and tissues undergoing EMT; second, occupancy of the FTS-1 site by these proteins in the chromatin of transitioning epithelia correlates with the activation of the EMT proteome; and third, the expression of recombinant *CBF-A* initiates key, early transcriptional and phenotypic events in EMT (see schematic in Figure 6).

The first component we found bound to FTS-1 was originally cloned as the CarG box–binding protein called CBF-A (26) — not

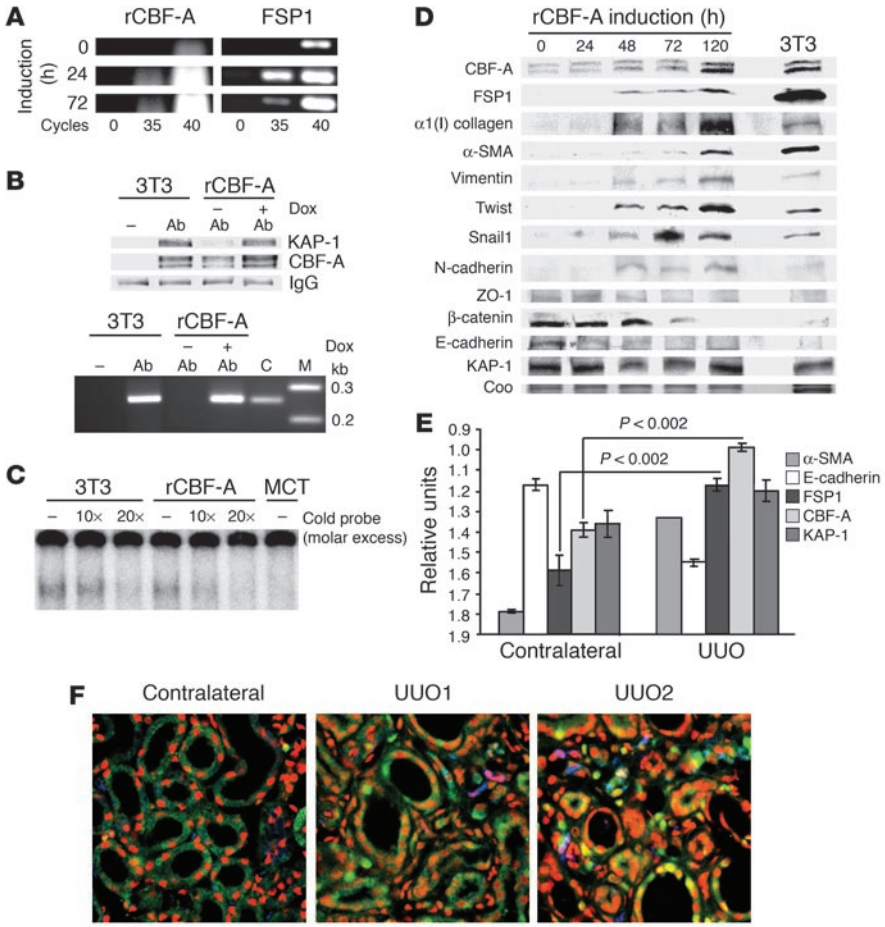


Figure 3

Overexpression of CBF-A in the epithelial cells leads to formation of the FTS-1 complex and expression of fibroblast markers. (**A**) Transcripts for *rCBF-A* and *FSP1* increase in PCR cycles after *rCBF-A* induction. (**B**) ChIP assay shows formation of the FTS-1 complex in epithelial cells expressing rCBF-A (+Dox; doxycycline) compared to 3T3 cells. IgG was a loading control. Immunoprecipitated DNA was used in PCR with primers for FTS-1. (**C**) Nuclear protein from MCT/rCBF-A$^+$ or 3T3 fibroblasts form FTS-1 complexes in EMSA with similar affinity, challenged by 10× or 20× molar excess of nonlabeled FTS-1 probe (Cold probe). MCT cells were the negative control. (**D**) Differential expression of epithelial and mesenchymal protein markers in MCT/rCBF-A$^+$ cells at different hours after induction of rCBF-A compared with 3T3 fibroblasts. Coo, Coomassie staining. (**E**) Increased expression of CBF-A and FSP1 in UUO undergoing EMT (28% and 26%, respectively) as measured by qRT-PCR and represented graphically in arbitrary units. Included for comparison are E-cadherin, α-SMA, and KAP-1. (**F**) CBF-A (yellow/orange) accumulates in the nuclei of tubular and interstitial cells from the UUO kidney in FSP1.GFP mice. FSP1$^+$ fibroblasts expressing GFP under the control of the *FSP1* promoter were recolored dark blue, and nuclei were counterstained red. Normal contralateral kidney tissue demonstrates CBF-A in the tubular cytoplasm (speckled green). In the cytoplasm, CBF-A and FSP1 staining merge to become light blue. Original magnification, ×630.

to be confused with other proteins named similarly: the CBF-binding intronic enhancer (29), Cbfa1/osteoblast-specific factor 2 (30), the core binding factors (CBFs) (31), PEBP2/CBFA1 (30), CCAAT-binding factor α (CBF-A) (32), and the hematopoietic transcription factor CBFA2 (33).

Multiple nucleic acid binding activities have been attributed to CBF-A, although its ability to bind a DNA duplex with sequence homology to FTS-1 has not been previously demonstrated. As a CArG box–binding protein, it was originally described as a transcriptional repressor of the α*SMA* promoter (26). There is no sequence similarity between FTS-1 and the canonical CarG box sequence (CC[AT]$_6$GG), and moreover, the SRE containing a canonical CArG box element does not compete the FTS-1 complex. CBF-A has also been described both as conditional repressor (34) or activator of the Ha-ras response element (35) and as an activator of the rat *spi 2* gene (36) via a GAGA box. CBF-A also binds an AT–rich promoter element in immunoglobulin *kappa* genes (37) and stabilizes a specific tetraplex telomeric DNA sequence (38). Thus, the overwhelming evidence implicates CBF-A as a transcriptional regulator and/or structural component of chromatin. However, remarkably, CBF-A contains no recognizable DNA binding motif for FTS-1.

CBF-A also belongs to the heterogeneous nuclear ribonucleoprotein (Hnrpab) subfamily A/B (hnRNP A/B). It contains 2 evolutionarily conserved RNA recognition motifs found in many RNA-binding proteins, functions as an mRNA-editing protein (39), and mediates the stability of proinflammatory mRNAs by binding to the AU-rich element (40). As a complex with actin, it is involved in nuclear-cytoplasmic transport of mRNA (41). Androgen-regulated posttranslational modification of CBF-A is also associated with cytoskeletal changes in a mouse fetal cell line, which may have been unrecognized EMT (42). Finally, CBF-A is upregulated in some mammary tumors (43) and in solid tumor metastasis (44). Depending on the context, CBF-A appears to be a "moonlighting" protein (45). It activates or represses transcriptional and posttranscriptional events, with a particularly important role in EMT. Finally, by inference, the association of EMT with metastatic conversion (13) together with reports of elevated levels of CBF-A in metastasis (43, 44) and its role in the transactivation of the Ha-ras protooncogene (35) suggests a possible role in carcinogenesis. These properties are completely consistent within our identification of CBF-A as a regular of EMT (13).

The second protein identified in the FTS-1 complex, KAP-1, is a well-studied transcriptional repressor that binds to KRAB domains on zinc finger proteins (27). The transcriptional repressor function of KAP-1 is mediated through its interaction with the heterochromatin-associated protein HP1 (46, 47) and with a specific histone methyltransferase, SETDB1, that contributes to the silencing of actively transcribed genes in euchromatin (48, 49). It is surprising, but intriguing, that the KAP-1 corepressor is present in our nucleoprotein complex that functions to activate transcription. Moreover, previous studies of KAP-1 have shown it to be almost exclusively bound to KRAB–zinc finger proteins in

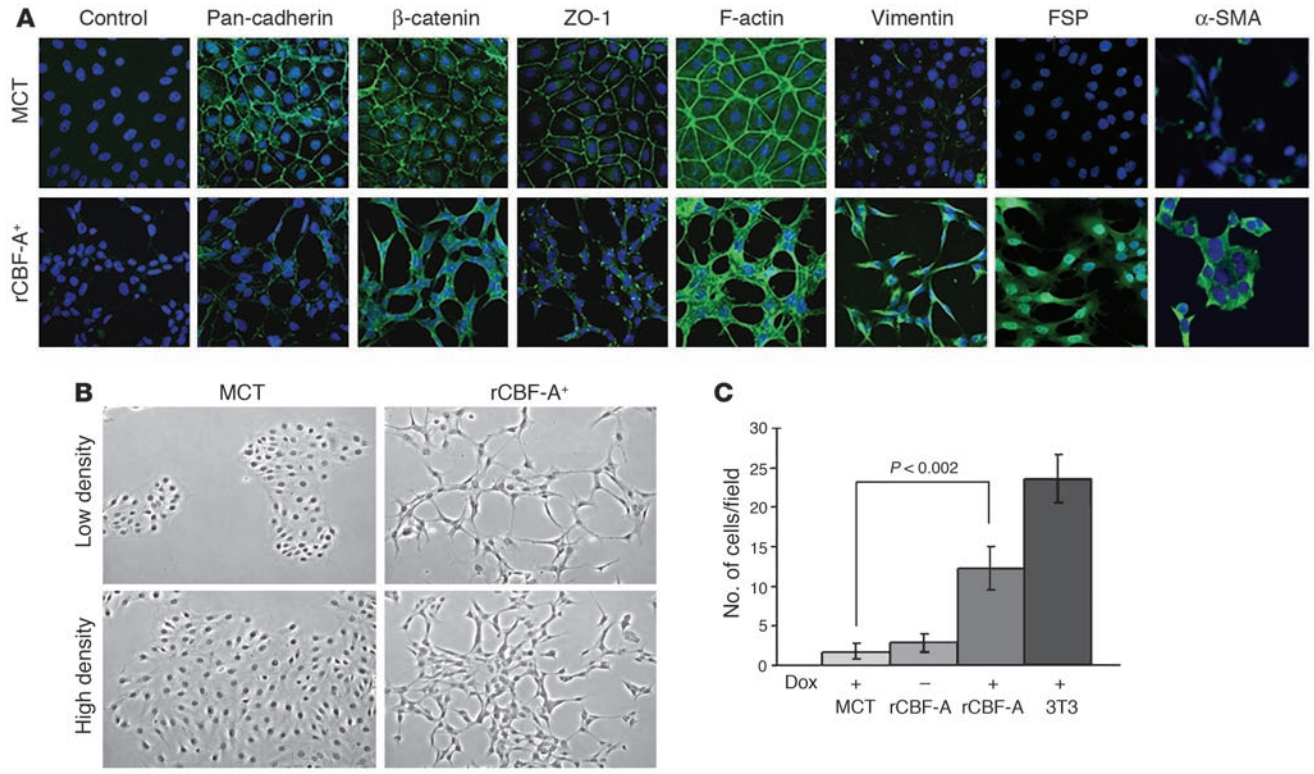


Figure 4
Inducible expression of *rCBF-A* in MCT epithelial cells leads to induction of FSP1, changes in the expression pattern of cell type markers, and acquisition of mesenchymal phenotype. (**A**) Immunostaining with cell type markers of naive epithelium (MCT) and of cells expressing rCBF-A. Results of staining with isotype IgG are included (Control). Original magnification, ×630. (**B**) Morphological changes in MCT/rCBF-A$^+$ compared with MCT cells at day 5 of *rCBF-A* induction. Cells are shown at both high and low density. Original magnification, ×200. (**C**) Increased motility of cells expressing *rCBF-A* (rCBF-A$^+$) after doxycycline induction as compared with rCBFA$^-$ and control MCT cells. 3T3 fibroblasts were included in this Boyden chamber assay as positive control. The average number of migrating cells is represented graphically and evaluated statistically.

vivo. However, our ChIP data conclusively show that both CBF-A and KAP-1 are present at the endogenous CBF-A site. There are a number of possibilities that may reconcile this new finding: first, there is evidence that the RNA recognition motif (RRM) in CBF-A may interact with the PHD domain of KAP-1 (50, 51). Thus, CBF-A and KAP-1 may form a physical complex that converts KAP-1 into a coactivator. We have recent evidence that posttranslational modification by SUMO regulates the repressor activity of KAP-1, thus providing a mechanism for toggling between activation and repression (unpublished observations). Second, and alternatively, an as-yet-unidentified KRAB–zinc finger protein may bind FTS-1 and tether KAP-1 to the site. Subsequently, the KAP-1 protein may remain associated with the chromatin after release of the putative KRAB–zinc finger protein during EMT activation. Third, CBF-A may be bound to a protein such as nucleophosmin, as has been recently described for the immunoglobulin *kappa* promoter (52). These findings open the possibility that the promoter-binding effect of CBF-A can be regulated by its interaction with various protein partners. The interaction of CBF-A with KAP1 may confer such specificity for binding the FTS-1 promoter element in epithelia transitioning to fibroblasts. How KAP-1 modulates this interaction is not yet known.

A key finding of this study is that *rCBF-A*-triggers expression of the EMT proteome in epithelia that parallels the changes in phenotype and potential for motility, all characteristics compatible with a fibroblast phenotype. Several other transcription proteins are also expressed during EMT. Twist (11), HMGA2 (21), and Snail (53) repress E-cadherin and induce EMT in epithelia, although loss of E-cadherin alone is not sufficient to initiate EMT (54). A recent report also suggests that gelsolin functions as a switch that controls conversion of E-cadherin to N-cadherin via Snail, and gelsolin's inhibition triggers EMT (55). Ets-1 (56) and Rho-related proteins (57) are also involved in this process. The fact that different transcriptional regulators can independently induce EMT suggests multifunctionality to molecular activation that may depend on the context of cytokine signaling events or changes in cytoskeletal structure that force cells into transition (1). Since metastatic tumor cells employ the same molecular program for EMT as that used by epithelia converting to fibroblasts (13), the FTS-1 complex is likely used by different epithelia to ensure phenotypic transitions.

Several potential binding sites for transcription factors have been identified in genes implicated in the EMT transcriptome, and our in silico search of the promoter regions of these genes upstream of transcriptional start site are summarized in Table 2. The E-box motifs are present only in 2 genes repressed in EMT, E-cadherin and gelsolin, but not in ZO-1 or β-catenin, where expression is equally attenuated. The CArG box motif is present in the repressed gene encoding E-cadherin but is also present

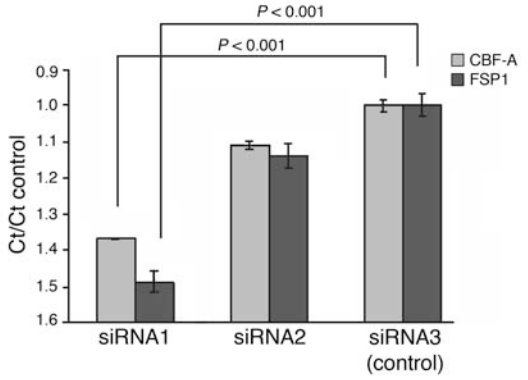


Figure 5
Inhibition of *CBF-A* transcripts by siRNA downregulates expression of *FSP1* transcripts. Three synthetic ribonucleotides were transfected into 3T3 fibroblasts, and the effect was monitored by qRT-PCR at different time points. The Ct values for transcripts encoding *CBF-A* and *FSP1* at the 48-hour time point obtained in triplicate were normalized to those for GAPDH. As siRNA3 did not attenuate the expression of either gene, the data were expressed as ratios to this reference control; siRNA1 best inhibited transcripts encoding *CBF-A* and *FSP1* compared with siRNA3; $P < 0.001$. Presented are combined results from 3 separate experiments.

in the activated gene encoding α-SMA. The lymphoid enhancing binding factor (LEF-1) binding site is also shared by up- and downregulated genes, reflecting perhaps the context-dependent effect of LEF-1. The most abundant regulation site is FTS-1, which again is evenly distributed among differentially expressed genes in EMT. It is likely that a specific combination of these sites is necessary to achieve differential expression in the EMT transcriptome. Accordingly, the promoter regions of *Twist*, *Snail*, *HMGA2*, *LEF-1*, and *Ets1* each contain FTS-1 sites, but not CArG- or E-box motifs (Table 2). Our findings implicate FTS-1 as a critical site that in combination with other *cis*-acting elements regulates the EMT transcriptome across a broad range of promoters.

In summary, this study reveals a mechanism for transcriptional activation of EMT using the interaction of FTS-1 sites with CBF-A and KAP-1. The finding that the CBF-A/KAP-1/FTS-1 complex is an activator of the genes encoding the EMT proteome suggests that it is an early proximal regulator, if not a candidate master gene, in the molecular program leading to fibroblast formation. This discovery has important implications for EMT-dependent biological processes, including organ fibrosis and metastatic cancer. CBF-A may be a good target candidate for suppressing EMT under pathologic circumstances.

Figure 6
This schematic, based on assumptions derived from the current data, shows CBF-A as a proximal transcription factor induced by outside-inside signaling. Migration to the nucleus causes the formation of CBF-A/KAP-1/FTS-1 binding complexes, which can activate other known transcriptional regulators of EMT, including Snail, Twist, HMGA2, LEF-1, and Ets-1. These are not the only EMT-related genes that are activated by CBF-A/KAP-1/FTS-1 binding complexes (see Table 2), but they could account for the activation of most of the known and well-studied transcriptional regulators of EMT. ILK, integrin-linked kinase.

Methods

Cell cultures. NIH 3T3 cells were used as a source of mouse fibroblasts. Cells of the kidney proximal tubular epithelial cell line MCT were derived from SJL/J mice (58). mIMCD cells were a gift from Raymond Harris (Vanderbilt University, Nashville, Tennessee, USA). Cell cultures were maintained as described previously (6). For induction of EMT in culture, epithelial cells were treated with 3 ng/ml of TGF-β and 10 ng/ml of EGF in standard culture media (6). The progress of EMT was monitored by phenotypic observation, cell motility assays, PCR, immunoblotting, or immunochemistry at different time points for up to 5 days after induction.

UUO. Three-month-old male BALB/c mice (The Jackson Laboratory) or FSP1.GFP mice on a BALB/c background were anesthetized according to our surgical protocol, and the right ureter was ligated at 2 points (4). Both fibrotic and normal contralateral kidneys were harvested 12 days after surgery and either snap-frozen in liquid nitrogen or fixed in 4% of paraformaldehyde. RNA was extracted from crushed frozen tissue by homogenization in TRIzol (Invitrogen), and after DNAse I treatment and inhibition (Ambion), 5-μg aliquots of RNA were used in a reverse transcription reaction with SuperScript reverse transcriptase (Invitrogen). The resulting cDNA was used as template for qRT-PCR analysis. Primers/probes for CBF-A, FSP1, E-cadherin, α-SMA, KAP-1, and GAPDH were obtained from TaqMan Gene Expression Assays (Applied Biosystems). Gene quantification was performed in triplicate on the Applied Biosystems 7900HT Fast Real-Time PCR System. Data were expressed in threshold cycle (Ct) values normalized against corresponding GAPDH expression. The Institutional Animal Care and Use Committee (IACUC) at Vanderbilt University approved all animal studies.

Immunohistochemistry, immunofluorescence, and immunoblotting. Polyclonal anti–KAP-1 antibody was prepared by Frank Rauscher (Wistar Institute, Philadelphia, Pennsylvania, USA); the antibody against CBF-A was kindly provided by Jonathan Dean, Imperial College of Science, London, United Kingdom; anti–pan-cadherin and anti–α-SMA antibodies were from

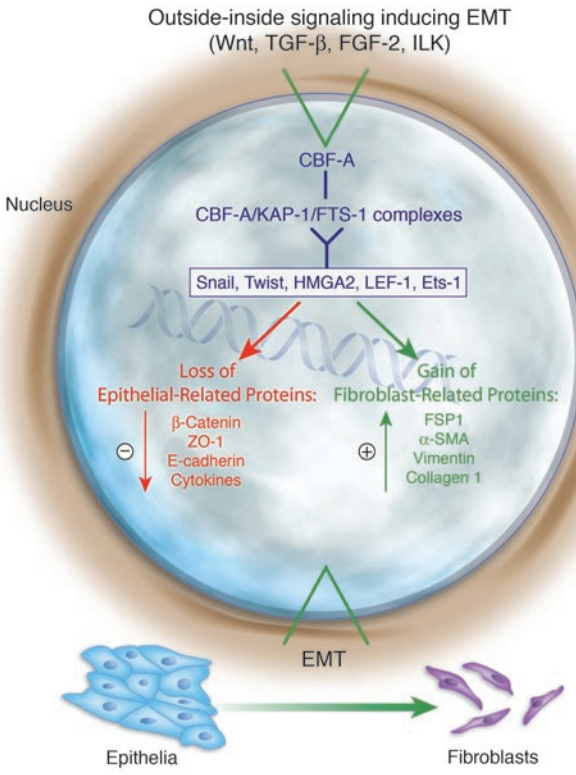


Table 2
Selected promoter sites in the EMT transcriptome

Genes	FTS-1	CArG	E-box	LEF-1
α-*Skeletal actin*	3	–	–	–
α-*SMA*	1	1	–	1
Annexin 8	1	–	–	–
β-*Catenin*	2	–	–	–
α*1(I) collagen*	–	–	–	1
α*2(I) collagen*	3	2	–	–
Desmoplakin	2	–	–	–
E-Cadherin	1	1	1	1
Ets-1	3	–	–	–
LEF-1	1	–	–	–
FSP1	2	–	–	–
Smad2, Smad3	–	–	–	–
Gelsolin	1	–	1	–
HMGA2	1	–	–	–
Laminin B3	2	–	–	1
Lamin A	3	–	–	–
MMP-9	2	–	–	–
Moesin	5	–	–	–
Rho	3	–	–	1
Snail	3	–	–	1
Slug	–	–	–	–
Twist	4	–	–	–
Vimentin	4	–	–	–
ZO-1	3	–	–	–

Sigma-Aldrich; antibodies against Collagen I, Vimentin, Twist, Snail-1, N-cadherin, ZO-1, β-catenin, and E-cadherin were from Santa Cruz Biotechnology Inc. Immunohistochemistry, immunofluorescence, and immunoblotting were performed as described previously (13, 59). Propidium iodide (0.1 μg/ml) (Molecular Probes; Invitrogen) and 1.0×10^{-8} M TRITC-conjugated phalloidin (Sigma-Aldrich) were used to stain nuclei and filamentous actins. Isotype variants were included as a negative control. Confocal fluorescence intensity was measured using the NIH Image program (http://rsb.info.nih.gov/nih-image). Image and data analysis were performed using the Vanderbilt University Medical Center Cell Imaging Core Resource.

ChIP. ChIP assays were performed using reagents from Upstate USA Inc. according to the manufacturer's protocol with the following modifications. Cross-linking by formaldehyde (1% vol/vol) was done directly in cultured fibroblasts or MCT/rCBF-A cells for 10 minutes at 37°C and terminated by glycine at a 0.125-M final concentration. The cells were washed twice with PBS containing proteinase inhibitors (1 mM PMSF, 1 mg/ml aprotinin, and 1 mg/ml pepstatin A). Nuclei were obtained using a nuclear purification kit (NUC-101; Sigma-Aldrich) supplemented with the same proteinase inhibitors according to the manufacturer's instructions. For each cell variant, 3 aliquots of 10^6 nuclei were resuspended in SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 7.9) plus protease inhibitor cocktail for mammalian tissues (Sigma-Aldrich). The nuclei were sonicated to yield DNA fragments ranging from 0.1 to 1 kb. Following centrifugation at 10,000 × *g* for 10 minutes at 4°C, one portion of the lysates was analyzed for DNA fragment size after reversing the cross-links and purification. The remaining supernatants were diluted 10-fold with ChIP dilution buffer and incubated by rotation with 80 μl of 50% salmon sperm DNA/protein A agarose beads for 90 minutes at 4°C. The beads were removed by brief centrifugation at 4°C, and 10 μl of either rabbit preimmune serum or rabbit CBF-A antibody was added to the lysates and incubated by rotation overnight at 4°C. The samples were subsequently incubated with 60 μl of protein A agarose beads for 1 hour at 4°C, and the beads were recovered by brief centrifugation and washed according to the manufacturer's protocol: once with low-salt 0.15 M NaCl buffer, once with high-salt 0.5 M NaCl buffer, and once with 0.25 M LiCl wash buffer. The pellets were washed once with 10 mM Tris-HCl and 1 mM EDTA, pH 8.0 (Tris-EDTA) and resuspended. The samples were split in 2, and the beads were collected by centrifugation; one portion was resuspended in SDS loading buffer, boiled for 10 minutes, and loaded on a ready-made 4%–20% SDS gradient gel (Bio-Rad) for Western blot analysis. The second portion for PCR analysis was resuspended in Tris-EDTA plus 0.2 M NaCl and incubated overnight at 65°C to reverse protein-DNA cross-links. This portion was used directly as a template for PCR analysis with primers producing an approximately 280-bp fragment encompassing the FTS-1 site in both directions; custom FTS-1 primers (Invitrogen) were: 5′-GGTGGTAGATATTCTGCTCC-3′ and 3′-TCTGCAACAACTCCTTGAGC-5′.

EMSA. Nuclei were purified from cultured cells as described above and either frozen or extracted. Nuclear extracts were prepared in 400 mM KCl buffered with 20 mM HEPES pH 7.9 supplemented with a cocktail of proteinase inhibitors (Sigma-Aldrich). The protein extracts were brought down to 100 mM KCl, and DNA binding activity was examined using the following oligonucleotide probes: FTS-1 (Invitrogen), regular: 5′CACTCACTACTTGATTGTGCCTGCT-3′, mutated: 5′-CACTCACTACCCAGCTGTGCCTGCT-3′; YB-1 consensus site (Santa Cruz Biotechnology Inc.), regular: 5′-AGACCGTACGTGATTGGTTAATCTCTT-3′, mutated: 5′-AGACCGTACGTGTAACCATAATCTCTT-3′; SRE (Santa Cruz Biotechnology Inc.), regular: 5′-GGATGTCCATATTAGGACATCT-3′, mutated: 5′-GGATGTCCATATTATTACATCT-3′. EMSA was performed as described previously (60). Briefly, 10–15 μg protein was incubated with 4×10^5 cpm of [^{32}P]ATP end-labeled oligonucleotides in 20 mM HEPES pH 7.9 with 100 mM KCl, 20% glycerol, and 0.2 mM EDTA. For competition analyses, protein was preincubated with unlabeled oligonucleotides or antibodies. FTS-1 complexes were resolved by 5% PAGE using 0.5× Tris-borate buffer pH 7.8, vacuum dried, and subjected to PhosphorImager analysis (GE HealthCare). Two-dimensional EMSA/SDS-PAGE gels were run with a nonradiolabeled probe; the position of the complex in the first dimension was deduced from radiolabeled duplicates, and the gel piece was cut and inserted on top of the SDS gel. As a negative control, nuclear extract without DNA was run in parallel in the first dimension (EMSA), and the area corresponding to the complex was cut and run together with the real complex in the subsequent second dimension (SDS-PAGE). For mass spectrometry analyses, complexes of unlabeled FTS-1 were identified by reversible silver staining (SilverQuest; Invitrogen) and matched to radiolabeled samples run in parallel. In each experiment, stained complexes were cut from the gel and subjected to trypsinization before analysis.

DNA affinity chromatography. Regular and mutated 100-bp biotinylated tetramers of the 25-bp FTS-1 oligonucleotide were custom made by Invitrogen and attached to Dynabeads M-280 Streptavidin (DYNAL; Invitrogen). FTS-1–specific complexes were formed with fibroblast nuclear extracts at the conditions described for EMSA. The nuclear extracts were incubated first with the mutated tetramer, and the flow-through was used for a secondary incubation with the normal FTS-1 tetramer. The beads were washed 3 times with incubation buffer, and fractions were eluted with 0.5 M KCl followed by the same buffer plus 5 M urea. The eluates were combined, dialyzed against 10 mM Tris-HCl pH 7.5, and analyzed by EMSA for their ability to form the FTS-1 complex. Positive fractions were trypsinized and subjected to tandem mass spectrometry for peptide identification.

Mass spectrometry analysis. Peptide mixtures were separated and analyzed using microcapillary reverse-phase HPLC (RP-HPLC) coupled to tandem

mass spectrometry (25). Peptide mixtures were loaded onto a 75-μm internal diameter RP-HPLC column (Poros R2; Perceptive Biosystems) equilibrated in 0.5% acetic acid and eluted using a linear gradient of 0%–40% acetonitrile over 60 minutes followed by 40%–60% over 10 minutes at a flow rate of 0.3 μl/min. Eluting peptides were analyzed by electrospray ionization mass spectrometry using an ion trap mass spectrometer (Finnigan LCQ Deca; Thermo Scientific). All tandem spectra were searched against the NCBI nonredundant protein database (http://www.ncbi.nlm.nih.gov/RefSeq/) using the SEQUEST algorithm (25, 61).

Generation of stable MCT/rCBF-A cell lines. A mammalian CBF-A/CBF-A expression vector, pSR-CBFA-N, was amplified using the Advantage PCR system (Clontech) with the following primers: the 5′ primer included an overhang with a unique *Bgl*II restriction site (ATTAGATCTGAGTAGCCGCTGCGGCCTGGCCGACCATGT) and the 3′ primer (ATTCTCGAGTCACTCGGCCGCTCCTGCTGCCTCTCAGTA) added a unique *Xho*I site. Following restriction digestion, the sequence was subcloned into a pcDNA5/TO vector from the T-REx tetracycline inducible system (Invitrogen), producing the pcDNA5/TO/rCBF-A vector. MCT cells were transfected simultaneously with both the pcDNA5/TO/rCBF-A vector and pcDNA6/TR regulatory vector of the T-REx system using SuperFect reagent (QIAGEN). Seven micrograms of TR plasmid and 3 μg of rCBF-A plasmid were used to insure higher copy number of the regulatory factor. Following transfection, cells were switched to media with tetracycline-free serum (Clontech). Twenty-four hours after transfection, recombinant cells were selected using 5 μg/ml blasticidin (Invitrogen). Once colonies formed, cells were tested for the presence of both vectors using the primers 5′-AGAGCTGCTTAATGAGGTCG-3′ and 5′-GCATTATATGCACTCAGCGC-3′ for the TR vector; and 5′-TGTCAGTGGAAGCAAGTGTG-3′ and 5′-ACTAGAAGGCACAGTCGAGG-3′ (vector BGH 3′-UTR specific) for rCBF-A. Cells were treated with 1 μg/ml doxycycline to induce rCBF-A and tested for changes in gene expression.

PCR analysis. All primers were custom made by Invitrogen: *FSP1*: sense, 5′-ATGGCAAGACCCTTGGAGGA-3′, anti-sense, 5′-CATTGCACATCATGGCAATG-3′; *GAPDH*: sense, 5′-GCAGTGGCAAAGTGGAGATT-3′, anti-sense, 5′-GCAGAAGGGGCGGAGATGAT-3′; *rCBF-A*: sense, 5′-TGTCAGTGGAAGCAAGTGTG-3′; anti-sense, 5′-ACTAGAAGGCACAGTCGAGG-3′ (derived from BGH 3′-UTR of the vector). Reverse transcription PCR was performed using 5 μg of total DNAse I–treated RNA with SuperScript reverse transcriptase (Invitrogen). Aliquots of cDNA, 2 μg or 0.2 μg, were subjected to PCR for 40 cycles, with 10-μl aliquots removed at 20, 25, 30, 35, and 40 cycles. Fragments were separated in agarose gels and stained with ethidium bromide, and the images were analyzed using Scion Image (Scion Corp.). For qRT-PCR analysis, total RNA was collected at different time points, treated with DNAse I, and converted by reverse transcription. Primers/probes for CBF-A, FSP1, and GAPDH were from TaqMan Gene Expression Assays (Applied Biosystems). Gene quantification was performed on the Applied Biosystems 7900HT Fast Real-Time PCR System.

siRNA inhibition. The Silencer Pre-designed siRNA oligonucleotides (Ambion) targeted exons 2 and 3 or exon 4 of CBF-A/CBF-A. These exons are not subject to alternative splicing. Each siRNA was separately transfected into 3T3 fibroblasts using Lipofectamine 2000 (Invitrogen) following the manufacturer's instruction for transfection. Negative control siRNA (Ambion) was transfected in parallel for comparison. The levels of CBF-A and FSP1 expression were monitored at different time points after transfection by qRT-PCR using extracted total RNA and primers/probes as described above. The Ct values obtained were normalized against the corresponding Ct values of GAPDH. The results from 3 separate experiments were examined for statistical differences. siRNA3 was selected as a negative control due to its lack of effect on CBF-A expression.

Motility assay. Migration through a membrane in response to a serum gradient (Boyden chamber assay) was performed in 24-well culture plates with 8-μm pore size inserts (BD). Cells were grown for 5 days with or without doxycycline. Cells (2×10^5) in 0.1% BSA-containing medium were loaded onto the top of each membrane. The bottom wells contained 15% FBS. When needed, doxycycline was added into both wells. The plates were incubated for 12 hours at 37°C; membranes were removed and stained with Diff-Quik stain (IMEB). Cells were monitored in each field by a digital camera using a ×10 objective lens (DM IRB; Leica). Six fields per membrane were counted in triplicate experiments.

Statistics. All values were expressed as mean ± SEM. Mann-Whitney U tests were used to compare the mean of non-parametric groups. Values of P less than 0.05 were considered statistically significant.

Acknowledgments

We thank Takeshi Miwa, Genome Information Research Center of Osaka University, for providing a mammalian CBF-A expression vector and Jonathan Dean, Imperial College of Science, London, for the anti–CBF-A antibody. We appreciate the valuable methodological help of Richard Caprioli and David Friedman at the Vanderbilt Mass Spectrometry facility. The authors are grateful to Linda Sealy (Vanderbilt) for critically reviewing an earlier version of the manuscript. E.G. Neilson was supported by NIH grants DK-46282, CA-68485, and DK-55926. A.J. Link was supported by NIH grants GM-64779, HL-68744, NS-43952, ES-11993, and CA-098131. Y.N. Dimitrova was supported by NIH grants GM-62112 and GM-75156 (Chazin).

Received for publication June 29, 2006, and accepted in revised form December 5, 2006.

Address correspondence to: Eric G. Neilson, Department of Medicine, Vanderbilt University School of Medicine, D-3100 MCN, Nashville, Tennessee 37232, USA. Phone: (615) 322-3146; Fax: (615) 343-9391; E-mail: eric.neilson@vanderbilt.edu.

1. Kalluri, R., and Neilson, E.G. 2003. Epithelial-mesenchymal transition and its implications for fibrosis. *J. Clin. Invest.* **112**:1776–1784. doi:10.1172/JCI200320530.
2. Thiery, J.P., and Sleeman, J.P. 2006. Complex networks orchestrate epithelial-mesenchymal transitions. *Nat. Rev. Mol. Cell Biol.* **7**:131–142.
3. Strutz, F., et al. 1995. Identification and characterization of a fibroblast marker: FSP1. *J. Cell Biol.* **130**:393–405.
4. Iwano, M., et al. 2002. Evidence that fibroblasts derive from epithelium during tissue fibrosis. *Invest.* **110**:341–350. doi:10.1172/JCI200215518.
5. Vongwiwatana, A., Tasanarong, A., Rayner, D.C., Melk, A., and Halloran, P.F. 2005. Epithelial to mesenchymal transition during late deterioration of human kidney transplants: the role of tubular cells in fibrogenesis. *Am. J. Transplant.* **5**:1367–1374.
6. Okada, H., Danoff, T.M., Kalluri, R., and Neilson, E.G. 1997. Early role of Fsp1 in epithelial-mesenchymal transformation. *Am. J. Physiol.* **273**:F563–F574.
7. Kawakita, T., et al. 2005. Intrastromal invasion by limbal epithelial cells is mediated by epithelial-mesenchymal transition activated by air exposure. *Am. J. Pathol.* **167**:381–393.
8. Nishitani, Y., et al. 2005. Fibroblast-specific protein 1 is a specific prognostic marker for renal survival in patients with IgAN. *Kidney Int.* **68**:1078–1085.
9. Lawson, W.E., et al. 2005. Characterization of fibroblast-specific protein 1 in pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **171**:899–907.
10. Jenkinson, S.R., Barraclough, R., West, C.R., and Rudland, P.S. 2004. S100A4 regulates cell motility and invasion in an in vitro model for breast cancer metastasis. *Br. J. Cancer.* **90**:253–262.
11. Yang, J., et al. 2004. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell.* **117**:927–939.
12. Sherbet, G.V., and Lakshmi, M.S. 1998. S100A4 (MTS1) calcium binding protein in cancer growth, invasion and metastasis. *Anticancer Res.* **18**:2415–2421.
13. Xue, C., Plieth, D., Venkov, C., Xu, C., and Neilson, E.G. 2003. The gatekeeper effect of epithelial-mesenchymal transition regulates the frequency of breast cancer metastasis. *Cancer Res.* **63**:3386–3394.
14. Grum-Schwensen, B., et al. 2005. Suppression of

tumor development and metastasis formation in mice lacking the S100A4(mts1) gene. *Cancer Res.* **65**:3772–3780.
15. Cho, Y.G., et al. 2005. Overexpression of S100A4 is closely associated with progression of colorectal cancer. *World J. Gastroenterol.* **11**:4852–4856.
16. Gupta, P.B., et al. 2005. The melanocyte differentiation program predisposes to metastasis after neoplastic transformation. *Nat. Genet.* **37**:1047–1054.
17. Moriyama-Kita, M., et al. 2005. S100A4 regulates E-cadherin expression in oral squamous cell carcinoma. *Cancer Lett.* **230**:211–218.
18. Shook, D., and Keller, R. 2003. Mechanisms, mechanics and function of epithelial-mesenchymal transitions in early development. *Mech. Dev.* **120**:1351–1383.
19. Kang, Y., and Massague, J. 2004. Epithelial-mesenchymal transitions; twist in development and metastasis. *Cell.* **118**:277–279.
20. Nieto, M.A. 2002. The snail superfamily of zinc-finger transcription factors. *Nat. Rev. Mol. Cell Biol.* **3**:155–166.
21. Thuault, S., et al. 2006. Transforming growth factor-beta employs HMGA2 to elicit epithelial-mesenchymal transition. *J. Cell Biol.* **174**:175–183.
22. Macias, D., Perez-Pomares, J.M., Garcia-Garrido, L., Carmona, R., and Munoz-Chapuli, R. 1998. Immunoreactivity of the ets-1 transcription factor correlates with areas of epithelial-mesenchymal transition in the developing avian heart. *Anat. Embryol.* **198**:307–315.
23. Lie-Venema, H., et al. 2003. Ets-1 and Ets-2 transcription factors are essential for normal coronary and myocardial development in chicken embryos. *Circ. Res.* **92**:749–756.
24. Okada, H., et al. 1998. Identification of a novel cis-acting element for fibroblast-specific transcription of the FSP1 gene. *Am. J. Physiol.* **275**:F306–F314.
25. Link, A.J., et al. 1999. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **17**:676–682.
26. Kamada, S., and Miwa, T. 1992. A protein binding to CArG box motifs and to single-stranded DNA functions as a transcriptional repressor. *Gene.* **119**:229–236.
27. Friedman, J.R., et al. 1996. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev.* **10**:2067–2078.
28. Kultz, D., Madhany, S., and Burg, M.B. 1998. Hyperosmolality causes growth arrest of murine kidney cells. Induction of GADD45 and GADD153 by osmosensing via stress-activated protein kinase 2. *J. Biol. Chem.* **273**:13645–13651.
29. Cohn, M.A., et al. 2001. Characterization of Sp1, AP-1, CBF and KRC binding sites and minisatellite DNA as functional elements of the metastasis-associated mts1/S100A4 gene intronic enhancer. *Nucleic Acids Res.* **29**:3335–3346.
30. Xiao, Z.S., Thomas, R., Hinson, T.K., and Quarles, L.D. 1998. Genomic structure and isoform expression of the mouse, rat and human Cbfa1/Osf2 transcription factor. *Gene.* **214**:187–197.
31. de Bruijn, M.F., and Speck, N.A. 2004. Core-binding factors in hematopoiesis and immune function. *Oncogene.* **23**:4238–4248.
32. Liang, S.G., and Maity, S.N. 1998. Pathway of complex formation between DNA and three subunits of CBF/NF-Y. Photocross-linking analysis of DNA-protein interaction and characterization of equilibrium steps of subunit interaction and dna binding. *J. Biol. Chem.* **273**:31590–31598.
33. Song, W.J., et al. 1999. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat. Genet.* **23**:166–175.
34. Gao, C., et al. 2004. S-nitrosylation of heterogeneous nuclear ribonucleoprotein A/B regulates osteopontin transcription in endotoxin-stimulated murine macrophages. *J. Biol. Chem.* **279**:11236–11243.
35. Mikheev, A.M., Mikheev, S.A., Zhang, Y., Aebersold, R., and Zarbl, H. 2000. CArG binding factor A (CBF-A) is involved in transcriptional regulation of the rat Ha-ras promoter. *Nucleic Acids Res.* **28**:3762–3770.
36. Leverrier, S., et al. 2000. Purification and cloning of type A/B hnRNP proteins involved in transcriptional activation from the Rat spi 2 gene GAGA box. *Biol. Chem.* **381**:1031–1040.
37. Bemark, M., Olsson, H., Heinegard, D., and Leanderson, T. 1998. Purification and characterization of a protein binding to the SP6 kappa promoter. A potential role for CArG-box binding factor-A in kappa transcription. *J. Biol. Chem.* **273**:18881–18890.
38. Weisman-Shomer, P., Cohen, E., and Fry, M. 2002. Distinct domains in the CArG-box binding factor A destabilize tetraplex forms of the fragile X expanded sequence d(CGG)n. *Nucleic Acids Res.* **30**:3672–3681.
39. Hou, V.C., et al. 2002. Decrease in hnRNP A/B expression during erythropoiesis mediates a pre-mRNA splicing switch. *EMBO J.* **21**:6195–6204.
40. Dean, J.L., et al. 2002. Identification of a novel AU-rich-element-binding protein which is related to AUF1. *Biochem. J.* **366**:709–719.
41. Percipalle, P., et al. 2002. Nuclear actin is associated with a specific subset of hnRNP A/B-type proteins. *Nucleic Acids Res.* **30**:1725–1734.
42. Umar, A., Luider, T.M., Berrevoets, C.A., Grootegoed, J.A., and Brinkmann, A.O. 2003. Proteomic analysis of androgen-regulated protein expression in a mouse fetal vas deferens cell line. *Endocrinology.* **144**:1147–1154.
43. Mikheev, A.M., et al. 2004. Frequent activation of CArG binding factor-A expression and binding in N-methyl-N-nitrosourea-induced rat mammary carcinomas. *Breast Cancer Res. Treat.* **88**:95–102.
44. Ramaswamy, S., Ross, K.N., Lander, E.S., and Golub, T.R. 2003. A molecular signature of metastasis in primary solid tumors. *Nat. Genet.* **33**:49–54.
45. Jeffery, C.J. 2003. Moonlighting proteins: old proteins learning new tricks. *Trends Genet.* **19**:415–417.
46. Ryan, R.F., et al. 1999. KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol. Cell. Biol.* **19**:4366–4378.
47. Lechner, M.S., Schultz, D.C., Negorev, D., Maul, G.G., and Rauscher, F.J., III. 2005. The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. *Biochem. Biophys. Res. Commun.* **331**:929–937.
48. Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., III. 2002. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* **16**:919–932.
49. Ayyanathan, K., et al. 2003. Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes Dev.* **17**:1855–1869.
50. Anderson, M., et al. 2002. A new family of cyclophilins with an RNA recognition motif that interact with members of the trx/MLL protein family in Drosophila and human cells. *Dev. Genes Evol.* **212**:107–113.
51. Fair, K., et al. 2001. Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. *Mol. Cell. Biol.* **21**:3589–3597.
52. Aranburu, A., Bennett, M., and Leanderson, T. 2006. The kappa promoter penta-decamer binding protein CBF-A interacts specifically with nucleophosmin in the nucleus only. *Mol. Immunol.* **43**:690–701.
53. Cano, A., et al. 2000. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* **2**:76–83.
54. Thiery, J.P. 2002. Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer.* **2**:442–454.
55. Tanaka, H., et al. 2005. siRNA gelsolin knockdown induces epithelial-mesenchymal transition with a cadherin switch in human mammary epithelial cells. *Int. J. Cancer.* **118**:1680–1691.
56. Gilles, C., Polette, M., Birembaut, P., Brunner, N., and Thompson, E.W. 1997. Expression of c-ets-1 mRNA is associated with an invasive, EMT-derived phenotype in breast carcinoma cell lines. *Clin. Exp. Metastasis.* **15**:519–526.
57. Savagner, P. 2001. Leaving the neighborhood: molecular mechanisms involved during epithelial-mesenchymal transition. *Bioessays.* **23**:912–923.
58. Haverty, T.P., et al. 1988. Characterization of a renal tubular epithelial cell line which secretes the autologous target antigen of autoimmune experimental interstitial nephritis. *J. Cell Biol.* **107**:1359–1368.
59. Bhowmick, N.A., et al. 2004. TGF-beta signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia. *Science.* **303**:848–851.
60. Venkov, C.D., Rankin, A.B., and Vaughan, D.E. 1996. Identification of authentic estrogen receptor in cultured endothelial cells. A potential mechanism for steroid hormone regulation of endothelial function. *Circulation.* **94**:727–733.
61. Eng, J., McCormack, A.L., and Yates, J.R., III. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Amer. Soc. Mass Spectrom.* **5**:976–989.

doi:10.1016/j.jmb.2007.03.047 J. Mol. Biol. (2007) 370, 269–289

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

The Structurally Disordered KRAB Repression Domain Is Incorporated into a Protease Resistant Core upon Binding to KAP-1-RBCC Domain

Hongzhuang Peng[1], Lisa C. Gibson[1], Allan D. Capili[2] Katherine L. B. Borden[3], Michael J. Osborne[3], Sandra L. Harper[1] David W. Speicher[1], Kehao Zhao[1], Ronen Marmorstein[1] Thomas A. Rock[1] and Frank J. Rauscher 3rd[1]*

[1]*The Wistar Institute, 3601 Spruce Street, Philadelphia PA 19104, USA*

[2]*Structural Biology Program Sloan-Kettering Institute New York, NY 10021, USA*

[3]*Institut de Recherche en Immunovirologie et Cancerologie, Universite de Montreal, 2950 Chemin Polytechnique, Pavillon Marcelle-Coutu, Montreal (Quebec) Canada H3T 1J4*

**Corresponding author*

The KRAB domain is a 75 amino acid transcriptional repression module that is encoded by more than 400 zinc finger protein genes, making it the most abundant repression domain in the human proteome. KRAB-mediated gene silencing requires a direct high affinity interaction with the RBCC domain of KAP-1 co-repressor. The structures of the free KRAB domain or the KRAB–RBCC complex are unknown. To address this, we have performed a systematic biophysical analysis of all KRAB isoforms using purified recombinant proteins. All KRAB domains are predominantly monomeric either alone or in a complex with KAP-1–RBCC protein, while a KRAB–SCAN isoform exists as a stable dimer. The KRAB:KAP-1–RBCC interaction requires only the A box in the context of the KRAB(Ab) or KRAB(AC) but both A and B boxes in the context of KRAB(AB). All isoforms bind the KAP-1–RBCC in a stable, zinc dependent fashion with a stoichiometry of KRAB1:3 RBCC with a zinc content of four atoms per RBCC monomer. Limited proteolysis, mass spectrometry and N-terminal sequence analyses suggest that a core complex comprises the entire RBCC domain of KAP-1 and the AB box of the KRAB domain rendering it resistant to proteolysis. NMR spectroscopy showed that unbound KRAB domain does not exist as a well-folded globular protein in solution but may fold into an ordered structure upon binding to the KAP-1–RBCC protein. This is the first example of a structurally disordered repressor domain that is the most widely conserved silencing domain in tetrapods.

Keywords: transcriptional repression; KRAB domain; KAP-1; protein–protein interaction; structurally disordered protein

Introduction

Biological processes such as transcriptional regulation, translation, and cellular signal transduction depend on the specific and temporal targeting of macromolecular complexes. The proteins involved in these processes are composed of multiple independently folded globular domains that cooperate in macromolecular recognition. These domains are usually recognizable by highly conserved signature sequence motifs and their presence and organization within a protein sometimes provides clues into the protein's function. Modular motifs within multi-domain proteins often specify specific interacting partners and the identification of these interactions has been beneficial in defining biochemical functions of particular proteins and their cellular pathways. Sequence homology prediction alone to

Abbreviations used: KRAB domain, Kruppel associated box; ZFP, zinc finger protein; KAP-1, KRAB associated protein 1; RBCC domain, RING finger, B boxes, and coiled-coil region; Ni-NTA, Ni^{2+}-intrilo-tri-acetic acid; DLS, dynamic light scattering; ICP, inductivity coupled plasma; HSQC, heteronuclear single quantum coherence.

E-mail address of the corresponding author: Rauscher@wistar.org

identify domain or/and motif type may lead to ambiguous assignments of domain type, making prediction of function a difficult task. Therefore, biochemical and biophysical characterization and determination of dynamics and the three-dimensional structure information are often essential for understanding biological function, especially for the domains that lack intrinsic globular structure under physiological conditions and fold into an ordered structure upon binding to their targets. With these principles in mind, we have been characterizing the Kruppel-associated box (KRAB) domain of the KRAB-zinc finger protein (KRAB–ZFP) family.

Modular transferable repression domains have emerged as a set of highly conserved structural motifs in families of transcription factors. These conserved repression domains include the BTB/POZ, KRAB, SCAN and SNAG modules. The BTB/POZ domain mediates homodimerization and heterodimerization between BTB/POZ proteins.[1,2] The structure of the BTB/POZ domain was determined and it exists as a homodimer.[3,4] The SCAN domain functions as a protein–protein interaction domain, mediating self-association or selective association with other proteins.[5,6] The isolated SCAN domain forms a stable dimer in solution,[7] and three-dimensional structure of this dimer from the human ZNF174 protein was recently determined.[8] We focused on the KRAB domain as a model system for the analysis of repression modules.[9–12] Analysis of the human genome sequence revealed 423 independent KRAB–ZFP genes, yielding alternative transcripts that altogether predict at least 742 structurally distinct proteins.[13] The KRAB domain is exclusively found in the NH_2 termini of Krupple-related ZFPs and it is classically divided into an A and a B box. KARB–ZFPs can be classified into several subfamilies: the first containing both a classical A and a classical B motif (KRAB(AB)), the second containing a classical A motif and a highly divergent b motif (KRAB(Ab)), and the third containing a classical A motif and another highly divergent C motif (KRAB(AC)) (Figure 1).[14] Some KRAB–ZFPs contain only an A box (KRAB(A)). In addition to the KRAB domain, some KRAB–ZFPs also carry a SCAN or leucine-rich (LeR) domain (Figure 1). Thus, the KRAB domain may represent structurally and functionally divergent properties relevant to its repression.

The KRAB domain was shown to be a potent, DNA binding-dependent transcriptional repression module.[10,15] The spatial and temporal expression of KRAB–ZFPs suggests that their biological functions could be involved in regulation of embryonic development, cell differentiation, and cell transformation.[16–19] Some KRAB–ZFPs are mainly restricted to lymphoid cells and may play a specific role in lymphoid differentiation, whereas others are expressed and specifically down-regulated during myeloid differentiation. Similarly, a recent study assigned a biological function in controlling sexual dimorphic liver gene expression to a set of KRAB–ZFPs (*Rsl1* and *Rsl2*).[20] A number of KRAB–ZFPs are candidate genes for human diseases based on their chromosomal locations.[21,22] Mutations of two KRAB–ZFP genes (ZNF41 and ZNF81) clustered on the X chromosome are found in patients with mental retardation, suggesting they are critical for development.[23,24]

Transcriptional repression by the KRAB domain correlates with its binding to KAP-1 (KRAB associated protein 1), also referred to as TIF1β or KRIP-1. KAP-1 belongs to a family of transcription co-regulators, which includs TIF1α, TIF1γ, TIF1δ, Bonus, and Ectodermin.[9,25–29] The TIF1 family encodes the signature RING-B box-coiled-coil tripartite motif (RBCC motif) at its NH_2 terminus, and a PHD-bromodomain at its COOH terminus. The RBCC motif likely functions as a cooperative protein/protein interaction motif.[12,30,31] We have characterized the KRAB:KAP-1 interaction extensively and this analysis has revealed that the RBCC functions as a protein interaction domain.[11,12] The KRAB domain binds directly with high affinity to the RBCC domain of KAP-1. This binding requires all three sub-domains of the RBCC domain and the mutation of any individual sub-domain abolishes binding. The KAP-1-RBCC domain oligomerizes as a homo-trimer, and oligomerization is obligate for KRAB binding. The stoichiometry of the KRAB:KAP-1 complex is 1:3. *In vivo*, KRAB-mediated gene silencing is absolutely dependent on KAP-1 binding.[9] The PXVXL motif of KAP-1 directly binds heterochromatin protein HP1.[32,33] The PHD and bromodomain of KAP-1 recruits the NURD histone deacetylase complex, SETDB1 histone H3-K9 methyl transferase, and other modification mechanisms involved in gene-silencing (A. Ivanov, unpublished data).[34–36] Although the KRAB–KAP-1-mediated silencing mechanism has been well characterized comparatively and the structural feature and properties of KRAB domain has been investigated,[37] the three-dimensional structure of KRAB domain remains to be determined.

It has been well established that the functions of many proteins are directly related to their three-dimensional structure. However, many proteins lack intrinsic globular structure under physiological conditions.[38] Such proteins are often involved in important regulatory functions within the cell and fold into an ordered structure upon binding to their target. An intrinsic lack of structure can confer functional advantages to a protein. There are numerous examples of domains, such as the kinase-inducible activation domain of the CREB and the acidic activation domain of p53 and VP16, that are unstructured in solution but become structured upon binding to their target proteins: CBP, MDM2, and TAFII31, respectively.[39–42]

We have performed a comprehensive and systematic analysis for the KRAB domains using biochemical, biophysical, and structural approaches. Our data indicate that the KRAB domain exists as a

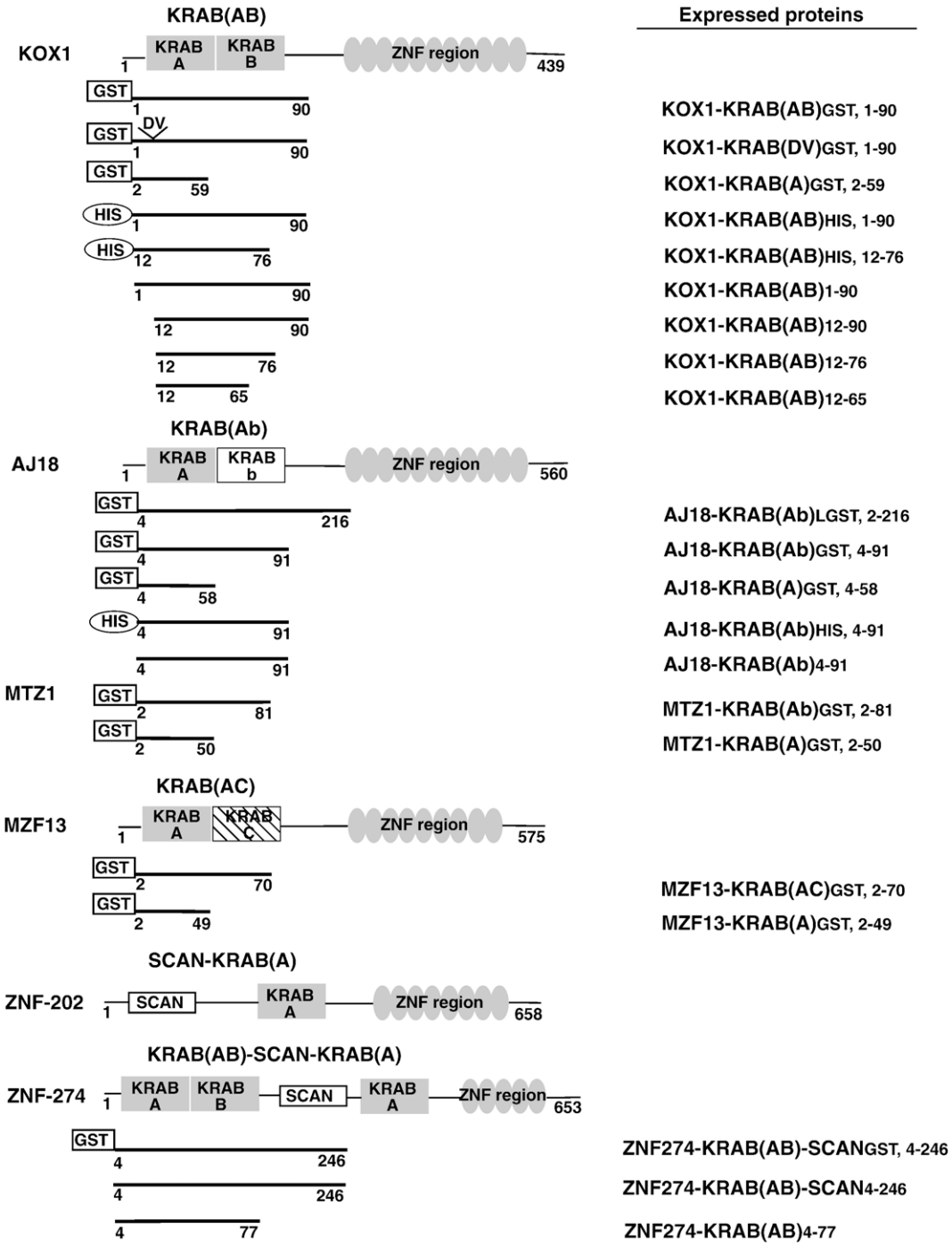


Figure 1. A diagram illustrating the architecture of KRAB–ZFPs and recombinant derivatives used in this study. The numbers represent amino acid position. A, A box; B, B box; b, b box; C, C box; ZNF region, zinc finger region; SCAN, SCAN domain. The column to the right indicates the expressed proteins. Database accession numbers: KOX1 (NM_015394); AJ18(NM_023988); MTZ1(L28167); MZF13(AF242376); ZNF202(NM_003455); ZNF274(NM_016325).

monomer. One molecule of KRAB binds three molecules of KAP-1–RBCC. We also show that the KRAB domain is a structurally disordered in solution and it actively binds to its ligand, the RBCC domain of KAP-1. These data may have broader implications for studying other transcriptional repressors and/or activators with their co-factors.

Results

Structure and sequence analyses of the KRAB-ZFPs family

KRAB–ZFP genes are very abundant in the human genome. We began our studies by aligning

and analyzing the protein sequence of the KRAB–ZFPs. The KRAB domain homology consists of ~75 amino acid residues and is predicted to fold into two amphipathic helices that are involved in protein–protein interaction (Figure 1).[43] The analyses indicate that KOX1–KRAB(AB) displays 54% identity (75% similarity) with AJ18–KRAB(Ab), and 55% identity (80% similarity) with MZF13–KRAB(AC), while AJ18–KRAB(Ab) displays 51% identity (82% similarity) with MZF13–KRAB(AC) over the whole KRAB domain. The high degree of identity and similarity among the subfamilies is exhibited only in the A box and no significant similarity was found between the other boxes (Figure 2(a)). The presence of B, b, or C boxes may influence the specific activities of different KRAB domains and the binding to the KAP-1 protein.[44] However, the structural basis of the B, b or C boxes in the recognition of their physiologically relevant ligand remains to be determined.

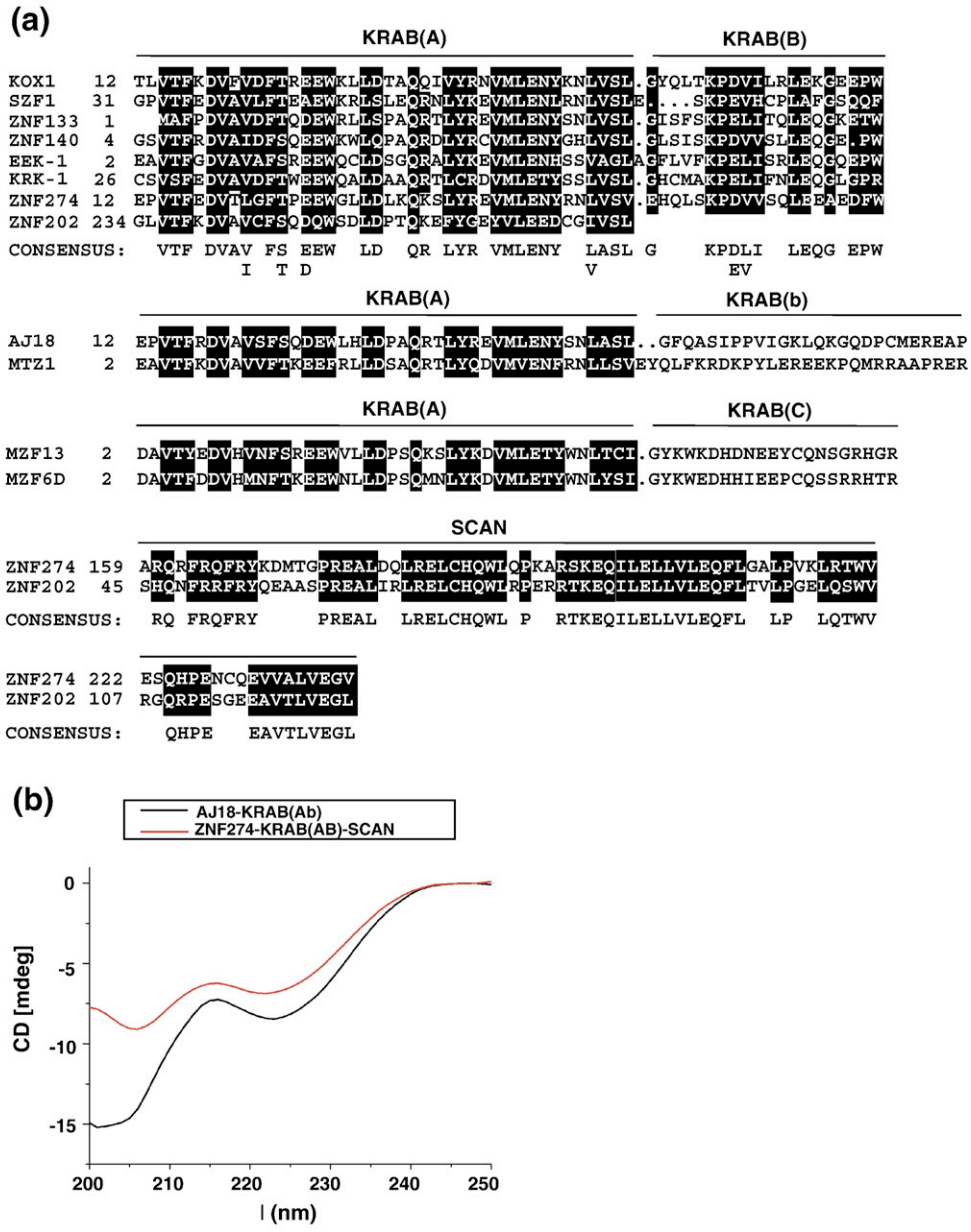


Figure 2. (a) Sequence alignment of the KRAB domain and SCAN domain from KRAB–ZFPs. Numbers refer to amino acid positions in the corresponding proteins. The KRAB and SCAN domains consensus residues are highlighted in black. The periods represent spaces introduced to obtain maximal alignment. (b) The CD spectra of the KRAB(Ab) of AJ18 and KRAB(AB)–SCAN of ZNF274. CD spectra were carried out in a buffer containing 20 mM Hepes (pH 7.5), 100 mM NaCl, 1 mM TCEP. A CD spectrum for a AJ18–KRAB(Ab) protein containing 65% random coil and 35% helix is plotted (black). A CD spectrum for a ZNF274–KRAB(AB)-SCAN protein containing 56% random coil, 39% helix, and 5% β sheet is plotted (red).

The program Predictprotein† suggests that the KRAB domain possesses a helical content of approximately 35% with two major helices and three loops. These helices have an amphipathic nature as defined by helical wheel analysis. These predictions indicate that the KRAB domain may provide an interface for protein–protein interactions and that it may bind to the RBCC domain of KAP-1 *via* a helix/helix interaction.

Expression and purification of recombinant KRAB domain proteins

The KRAB domains utilized here include seven members from KRAB(AB), two from KRAB(Ab), one from KRAB(AC), one from SCAN-KRAB(A), and one from KRAB(AB)-SCAN–KRAB(A) subfamilies (Figure 1; Table 1). We selected the KRAB domains from all subfamilies for the following reasons: (1) the KOX1-KRAB(AB) domain was originally utilized to isolate the KAP-1 co-repressor, and mutations in this domain that concomitantly abolish repression and KAP-1 binding are well characterized;[9] (2) the KOX1–KRAB(AB) domain is highly expressed in *Escherichia coli*, and is well behaved in protein reconstitution assays;[11,12] (3) the GAL4-KOX1–KRAB(AB) and GAL4-SZF1–KRAB (AB) fusion is a potent, KAP-1-dependent, and DNA binding-dependent transcriptional repressor *in vivo*;[10,45] (4) we hypothesized that the variant amino acids in B, b, and C boxes may influence the folding of the KRAB domain so that the solubility and behavior of the proteins may be different.

To optimize the protein expression level and solubility, three protein expression systems were used in *E. coli*. A baculovirus expression system was utilized as well in the event that post-translation modification might become important (Table 1). All His-tagged proteins were first purified using Ni^{2+}-NTA chromatography under native conditions, followed by gel filtration chromatography. Most of the His-tagged KRAB(AB) domain proteins were not soluble under native conditions except His-tagged AJ18-KRAB(Ab) (Table 1). The GST-fusion proteins were purified with GST resin, followed by a thrombin digestion to release the target proteins. Further purification was carried out using ion-exchange and/or gel filtration chromatography (data not shown). GST-AJ18–KRAB(Ab) and GST-ZNF274–KRAB(AB) were highly soluble and stable under native conditions. SDS–PAGE analyses revealed that the purified soluble His-tagged or GST-fusion proteins migrated as predicted by their molecular mass. Production of these highly soluble proteins allowed us to conduct further functional and structural analyses.

To address whether the highly soluble KRAB domain proteins were structured, dynamic light scattering (DLS) analyses were performed using a DynaPro-801 molecular sizing instrument. These analyses showed that most of the KRAB proteins display moderate polydispersity, with the AJ18–KRAB(Ab) protein exhibiting the greatest polydispersity, indicating that they are generally prone to aggregation. CD spectra of AJ18–KRAB(Ab), and ZNF274–KRAB(AB)-SCAN proteins showed that they contain mostly random coil (56% ~ 65%) with a small degree of helix (35% ~ 39%) but little β sheet (0 ~ 5%) (Figure 2(b)), consistent with the secondary structure predicted by Predictprotein. These data suggest that the KRAB domain alone may not be highly structured.

The KRAB domain of ZFP exists in a monomeric state in solution

We previously employed biochemical and biophysical analyses for the KOX1–KRAB(AB):KAP-1–RBCC complex and its components. The KOX1–KRAB(AB) was a soluble self-aggregate under physiological buffer conditions. The non-aggregated portion of KOX1–KRAB(AB) could be separated from aggregate by gel filtration in the presence of non-denaturing detergent (Figure 3(a), I). Although it is predominantly monomeric in the absence of KAP-1, this non-aggregated KOX1–KRAB(AB) protein is highly active for binding to the KAP-1–RBCC when these factors co-exist in solution. To explore the oligomeric state of other KRAB domains, gel filtration chromatography and analytical ultracentrifugation were used. Gel filtration studies showed that under physiological buffer conditions, the AJ18–KRAB(Ab) protein (calculated M_r of monomer is 10 kDa) eluted in a single peak = 17 kDa, consistent with a monomer (Figure 3(a), II). In contrast, the ZNF274–KRAB(AB)-SCAN protein (calculated M_r of dimer is 57.41 kDa) eluted in a single peak = 130 kDa, consistent with a tetramer or an elongated dimer (Figure 3(a), III).

To more rigorously establish the oligomeric state of the KRAB domains, we employed equilibrium sedimentation experiments. Analyses were performed at 4 °C at different speeds and protein concentrations. The concentration of protein *versus* radius data was fitted with various models of self-association using non-linear regression.[46] For AJ18–KRAB(Ab) protein, the data were best described by a model containing predominantly monomer (90% ~ 95%) and a small amount of dimer with K_d value of 1.3 mM at speed of 38,900 rpm, and with K_d value of 5.1 mM at speed of 55,000 rpm (Figure 3(b)), indicating that the AJ18–KRAB(Ab) protein is predominantly a monomer. For the ZNF274–KRAB (AB)-SCAN protein, the data were best fit by a model containing a single species of dimer with a M_r of 57,701 kDa at speed of 23,200 rpm (Figure 3(c)). This value is close to the calculated M_r of the dimer 57,412. This result indicates that the ZNF274–KRAB (AB)-SCAN protein is a dimer and suggests that the 130 kDa size observed from gel filtration is due to an asymmetric shape. The data also suggests that the SCAN domain is responsible for dimerization and is consistent with the observation that the SCAN domain from ZNF174 also forms a dimer.[8]

† http://www.cubic.bioc.columbia.edu/predictprotein

Table 1. Expression, solubility, and function for the KRAB domain of KRAB–ZNPs and the RBCC domain of KAP-1 protein

Expression system tag	Proteins	Solubility[a]	KAP-1-binding[b]
KRAB domain			
Intein	KOX1-KRABIntein, 1–90	+	+
6Histidine	KOX1-KRAB(AB)HIS, 1–161	–	++
	KOX1-KRAB(AB)HIS, 1–90	–	++
	KOX1-KRAB(AB)HIS, 12–76	–	++
	SZF1-KRAB(AB)HIS, 3–110	–	++
	ZNF133-KRAB(AB)N-HIS, 1–119	–	ND
	ZNF133-KRAB(AB)C-HIS, 1–119	–	ND
	ZNF140-KRAB(AB)HIS, 1–119	–	ND
	ZNF141-KRAB(AB)HIS, 1–107	–	ND
	EEK1-KRAB(AB)HIS, 1–122	–	ND
	KRK1-KRAB(AB)HIS, 16–127	–	ND
	AJ18-KRAB(Ab)HIS, 4–91[d]	+++	++
GST	KOX1-KRAB(AB)GST, 1–90	+	++
	KOX1-KRAB(DV)GST, 1–90	+	–
	KOX1-KRAB(A)GST, 2–59	+	–
	SZF1-KRAB(AB)GST, 3–110	+	++
	AJ18-KRAB(Ab)LGST, 4–216	+	ND
	AJ18-KRAB(Ab)GST, 4–91[c,d]	(+++)	(++)
	AJ18-KRAB(A)GST, 4–58	+++	++
	MTZ1-KRAB(Ab)GST, 2–81	+	++
	MTZ1-KRAB(A)GST, 2–50	+	++
	MZF13-KRAB(AC)GST, 2–70	+	++
	MZF13-KRAB(A)GST, 2–49	+	++
	ZNF202-SCAN-KRAB(A)GST, 37–286	+	–
	ZNF274-KRAB(AB)-SCANGST, 4–246[c]	(+++)	(+)
	ZNF274-KRAB(AB)GST, 4–77[d]	(+++)	(++)
6Histidine baculovirus	ZNF140-KRAB(AB)HIS, 1–119	–	ND
RBCC domain			
6Histidine	KAP-1-RBCCHIS, 22–418	++	++
	KAP-1-RBCCHIS, 50–418	++	++
	KAP-1-RBCCHIS, TEV, 50–418	++	++
	KAP-1-B1B2CCHIS, 136–418	–	–
	KAP-1-B2CCHIS, 199–418	+	–
6Histidine baculovirus	KAP-1-RBCCHIS, 22–418	++	++
pRM1 and 6histidine	KOX1-KRAB(AB), 1–90[c]	+++	++
	KAP-1-RBCCHIS, 22–418	+++	++
	KOX1-KRAB(AB), 12–76[c]	+++	++
	KAP-1-RBCCHIS, 22–418	+++	++
	KOX1-KRAB(AB), 12–90[c]	+++	++
	KAP-1-RBCCHIS, 50–418	+++	++
	KOX1-KRAB(AB), 12–65[c]	+++	++
	KAP-1-RBCCHIS, 50–418	+++	++
	KOX1-KRAB(AB), 12–90[c]	+++	++
	KAP-1-RBCCHIS, TEV, 50–418	+++	++
	AJ18-KRAB(Ab), 4–91	+++	++
	KAP-1-RBCCHIS, 50–418	+++	++
6 Histidine baculovirus	ZNF140-KRAB(AB)HIS, 1–119	–	ND
	KAP-1-RBCCHIS, 22–418	++	ND

ND, not determined. Parentheses indicate that the protein was tested after cleavage of the GST tag.

[a] Solubility indicates the initial solubility of the expressed proteins. A– means that the proteins are not soluble under native purification conditions. These proteins were purified under denaturing conditions and were refolded by either step dialysis or on a column for functional studies. A+ indicates that the proteins are soluble under native purification conditions.

[b] KAP-1 binding indicates the KRAB domain proteins were tested by GST-association assays, competition assays, transient transfection assays for repression.

[c] The proteins or protein complexes were set for crystallization screen using screen kits including (1) Crystal Screen, (2) Crystal Screen 2, (3) Wizard I, (4) Wizard II, and (5) Index at temperature 25 °C and 4 °C.

[d] The proteins were studied by NMR spectroscopy.

The KRAB(AB), KRAB(Ab), KRAB(AC) domains form a complex with the RBCC domain of KAP-1

We reconstituted the KOX1–KRAB(AB):KAP-1–RBCC protein complex using recombinant proteins *in vitro*.[11,12] The yeast two-hybrid assay indicates that the KRAB(AB), KRAB (Ab) and KRAB (AC) interact with KAP-1 *in vivo*, and that the B, b and C boxes influence this interaction differently.[44] To define which of the A, B, b, and C boxes of KRAB are integral to its interaction with the KAP-1–RBCC, we prepared a series of B, b, or C box deletions in each of the KRAB domains in GST-fusion format (Figure 1; Table 1). The GST-fusion proteins were

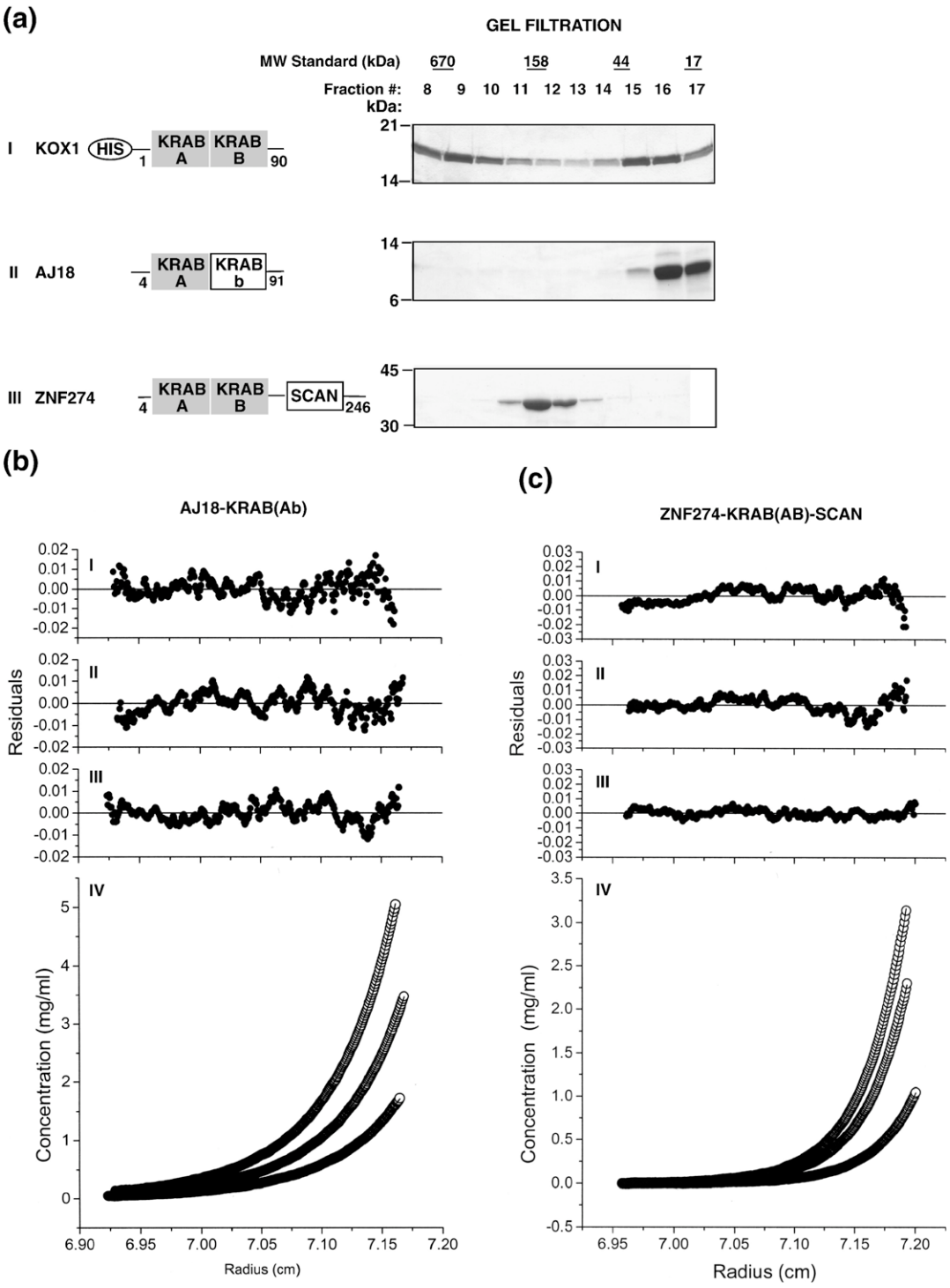


Figure 3. (a) Gel filtration analyses of KOX1–KRAB(AB), AJ18–KRAB(Ab), and ZNF274–KRAB(AB)-SCAN proteins. The proteins were resolved on a Superdex 200 column. Fractions were collected and proteins were analyzed by SDS–PAGE. The elution profiles of KOX1–KRAB(AB) protein (I), AJ18–KRAB(Ab) protein (II), and ZNF274–KRAB(AB)-SCAN protein (III) on the gel filtration column. (b) Representation of analytical ultracentrifugation analysis of AJ18–KRAB(Ab). Sedimentation equilibrium analyses were performed with three different loading concentrations in separate cells at 38,900 rpm and 55,000 rpm at 4 °C. The circles in the main panel (IV) show the concentration *versus* radius data for the three cells at equilibrium. The three data sets were fitted with a model describing monomers. I–III shows the residuals for the data points to the fitted curves at the three protein concentrations displayed from highest concentration (top) to lowest concentration. (c) Representation of analytical ultracentrifugation analysis of ZNF274–KRAB(AB)-SCAN. The circles in the main panel (IV) show the concentration *versus* radius data for the three cells at equilibrium. The three data sets were fitted with a model describing dimers. I–III shows the residuals for the data points to the fitted curves at the three protein concentration displayed from highest concentration (top) to lowest concentration.

expressed and purified in *E. coli*, and analyzed for binding to the KAP-1–RBCC by GST association assay (Figure 4(a)). These studies reveal that deletion of the B box in KRAB(AB) exhibits greatly diminished binding to KAP-1–RBCC. A similar result was obtained from the KRAB-O protein in which the deletion of B box significantly reduced KAP-1–RBCC association (data not shown). The deletion of B box has the same effect as the KRAB (DV-AA) mutant, which has been shown to abolish KAP-1 binding and repression function.[10] This result suggests that the B box is necessary for KRAB(AB) binding to the RBCC domain and is consistent with previously published *in vivo*

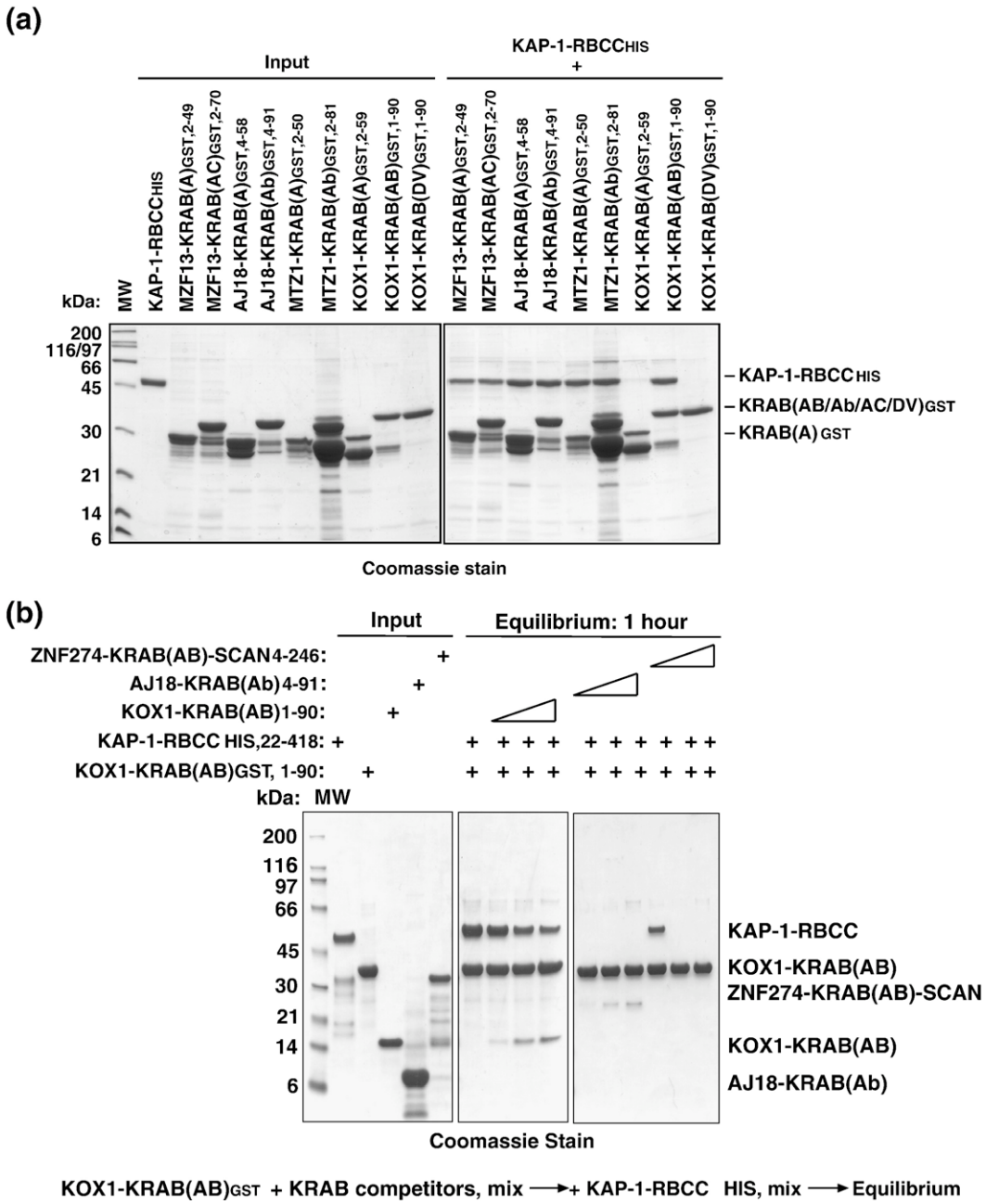


Figure 4. (a) The binding of the KAP-1–RBCC to KRAB(AB), KRAB(Ab), KRAB(AC) but not mutant form as detected by GST association assay. The input lanes represent 5 μg of purified KAP-1–RBCC protein that was added to each binding reaction mixture. No binding was detected for the GST–KOX1–KRAB(A) and GST–KOX1–KRAB(DV-AA). Deletion of the B box of KOX1–KRAB has the same effect as the KRAB(DV-AA) mutant. (b) Competition of the GST–KOX1–KRAB(AB): KAP-1-RBCC complex formation by the purified KOX1–KRAB(AB), AJ18–KRAB(Ab), and ZNF274–KRAB(AB)-SCAN proteins. A constant amount of GST–KOX1–KRAB(AB) (5 μg) was mixed with increasing amounts of one of the competitors (1 μg, 5 μg, and 10 μg), and then a constant amount of KAP-1–RBCC (5 μg) was added. The reaction was incubated at equilibrium conditions for one hour at room temperature. AJ18–KRAB(Ab) and ZNF274–KRAB(AB)-SCAN are very efficient competitors of the GST–KOX1-KRAB(AB):KAP-1–RBCC interaction.

studies.[44] However, deletion of the b box or C box in KRAB(Ab) or KRAB(AC) had no effect on their binding to the RBCC domain, indicating that the A box in the context of KRAB(Ab) and KRAB (AC) is sufficient to bind the RBCC domain *in vitro* and is consistent with previously published *in vivo* studies.[44] Together, the above data indicate that the A box in KRAB(AB) might fold differently in KRAB(Ab) or KRAB(AC); and that the B box in KRAB(AB) is needed for recognition of the RBCC domain, possibly for proper folding of A box of KRAB(AB).

To determine whether the KRAB(AB) and KRAB (Ab) domain have the same binding affinity to the RBCC domain, we performed a competition assay under equilibrium conditions. The GST-KOX1–KRAB(AB) was first immobilized on GST resin, and increasing amounts of competitors including KOX1–KRAB(AB), ZNF274–KRAB(AB)-SCAN, and AJ18–KRAB(Ab) were added to the resin, followed by addition of a constant amount of KAP-1–RBCC protein. The reaction was then incubated under equilibrium conditions. The results show that ZNF274–KRAB(AB)-SCAN and AJ18–KRAB(Ab) are very efficient and comparable competitors for the GST-KOX1–KRAB(AB):KAP-1–RBCC interaction (Figure 4(b)).

Stoichiometry of the RBCC domain of KAP-1 to zinc is 1:4

We previously defined the RBCC region to be sufficient and necessary for oligomerization of KAP-1 and KRAB domain binding using recombinant proteins *in vitro*,[12] although two B boxes and the coiled-coil region were shown to be sufficient for interaction with the KRAB domain using a *in vivo* mammalian two-hybrid assay.[47] Obviously, the KAP-1–RBCC mediates a highly specific, direct interaction with KRAB domain, and it functions as an integrated, cooperative structural unit wherein each sub-domain contributes to oligomerization and ligand recognition. The RING finger is a cysteine-rich motif of the form C3HC4 and binds two molecules of zinc in a unique cross-braced ligation system.[48–50] The RING finger is likely to contribute either to specificity and/or multimerization properties of the RBCC motif. RING finger structures suggest that a common hydrophobic core is formed as a result of zinc-chelation, and that the variable sequence between the conserved ligation residues provides specificity for the protein recognition. The B-box is also a cysteine-rich zinc binding motif in the form of CHC3H2.[51] The NMR structure of the XNF7 B-box shows that only one zinc atom is bound with the other potential chelation residues unoccupied. The coiled-coil motif is comprised of appropriately spaced hydrophobic residues predicted to form an extended α-helical region. The coiled-coil region of KAP-1 is predicted to fold into two potential leucine zipper-like motifs,[12,52] where each region is likely to be an amphipathic α-helix.

The conservation of eight potential metal ligands in the RING finger and B-box in the RBCC domain suggests that it is a zinc-binding protein. Mutations of these zinc-chelating cysteine and histidine interrupt the folding of the molecule and abolish KRAB domain binding.[12] To determine the identity and stoichiometry of metal bound by the RBCC domain, we used inductivity coupled plasma (ICP) spectrometry. These measurements indicate that 330 μM KAP-1–RBCC protein contains 1. 54 mM zinc, and 33 μM KAP-1–RBCC protein contains 143 μM zinc. The ratio of one protein molecule to zinc atoms is 1:4.6 and 1:4.29, respectively, indicating that RBCC domain binds four zinc atoms per protein molecule. The negative control of His-tagged protein shows very low background binding to zinc atoms (data not shown). Potentially, the RBCC domain utilizes all eight conserved cysteine and histidine residues from the RING finger for metal ligation. Only four conserved cysteine and histidine residues from the B1 and B2 boxes are used for metal chelation. This result is consistent with observations from the NMR structure of the XNF7 B-box.[53] Together, the mutation study[12] and zinc content analyses indicate that the RBCC domain requires zinc for folding.

Reconstitution of KRAB:KAP-1–RBCC protein complex *in vivo* using a two-plasmid expression system

We previously mapped the interaction domains for both KOX1–KRAB(AB) and KAP-1–RBCC proteins *in vivo* and *in vitro* and reconstituted the protein complex using these highly purified proteins. However, the efficiency of forming this protein complex is not high enough to allow us to do biophysical and structural studies. Therefore, we developed a two-plasmid expression system to co-express and co-purify the protein complex. This system increases protein expression, solubility, efficiency of protein complex formation, and overall complex stability. As a result, we successfully obtained a highly purified and stable protein complex (Figure 5(a)). The stoichiometry of KRAB:KAP-1–RBCC was then determined by sedimentation equilibrium analysis. The datum followed the model predicting a 1:3 molar ratio of KRAB(AB):KAP-1–RBCC complex (79%, 81%, and 92% in three protein concentrations, respectively) and a small amount of aggregate (21%, 19%, and 8%, respectively) at a speed of 10,000 rpm (Figure 5(b)). This result was consistent with previous observations from densitometry and kinetic studies of the KRAB (AB):KAP-1–RBCC protein complex.[11]

The entire RBCC domain of KAP-1 and the AB box of KRAB forms a core complex

To determine the core complex of the KRAB:KAP-1–RBCC proteins, we employed limited proteolysis experiment, where the core would be protected from the protease digestion. Proteases including trypsin, proteinase K, subtilisin, and thermolysin were used

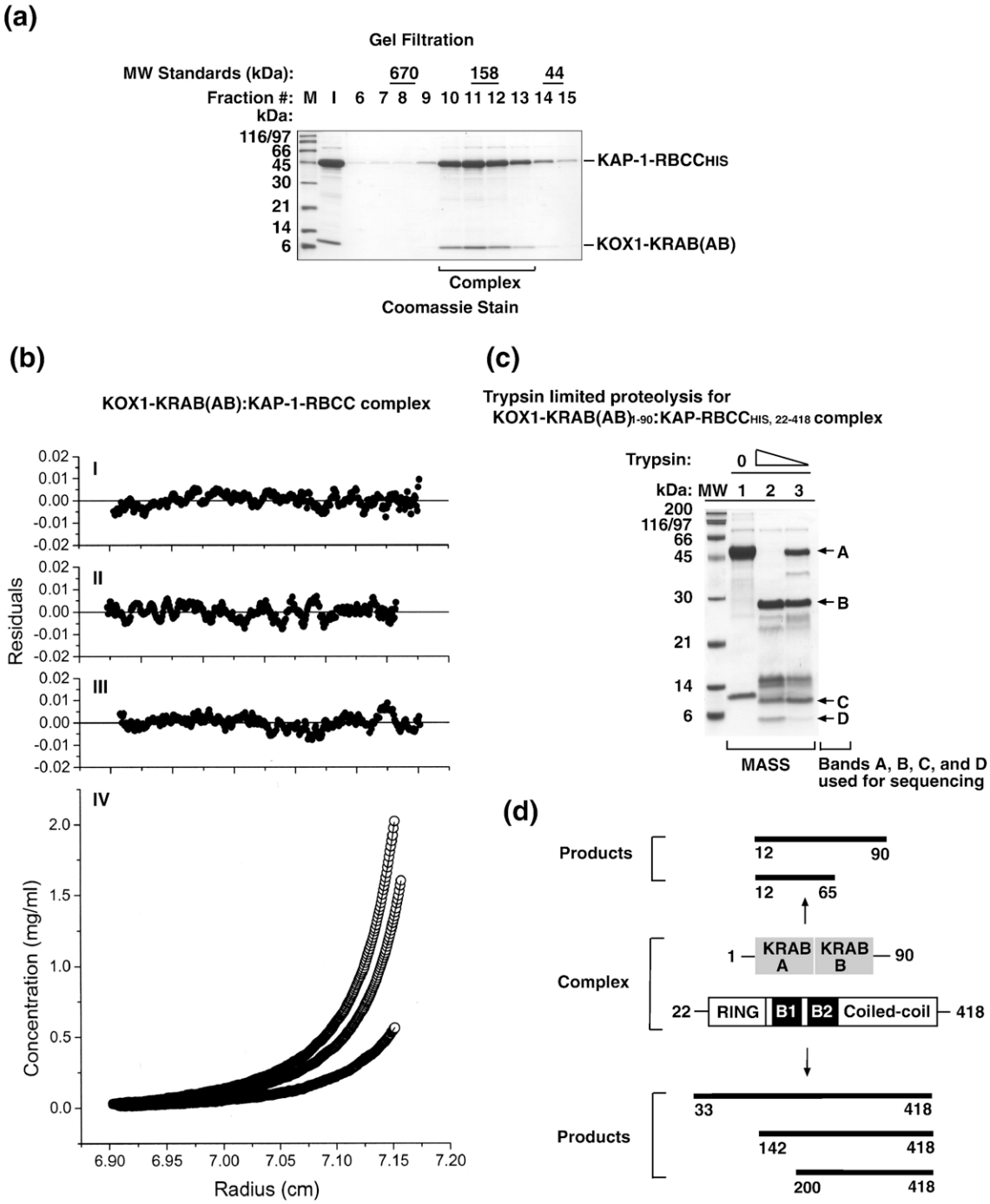


Figure 5. (a) Gel filtration analysis of the KOX1–KRAB(AB):KAP-1–RBCC protein complex. The reconstitution of KOX1–KRAB(AB):KAP-1–RBCC protein complex *in vivo* is described in Materials and Methods. The protein complex was resolved on a Superdex 200 column. The fractions were collected, and the proteins were analyzed by SDS–PAGE. The peak fraction of the complex (fraction 11) eluted with a molecular mass of ~158 kDa. (b) Analytical ultracentrifugation analysis of KOX1–KRAB(AB):KAP-1–RBCC protein complex. Sedimentation equilibrium analysis was performed at 10,000 rpm and 4 °C. Panel IV shows the concentration *versus* radius data for three loading concentration of the protein complex (circles). The three data sets were fitted globally with a model describing a predominant complex with 1:3 molar ratio of KOX1–KRAB(AB):KAP-1–RBCC and small amount of aggregate. I–III shows the residuals for the data points to the fitted curves at the three protein concentrations displayed from highest concentration (top) to lowest concentration. (c) Trypsin limited proteolysis, Mass spectrometry, and N-terminal sequence analyses for KOX1–KRAB(AB)1–90:KAP–RBCCHIS,22–418 protein complex. The protein complex was treated with trypsin (0, lane 1; 0.04 μg/μl, lane 2; and 0.008 μg/μl, lane 3). Aliquots of each reaction were analyzed by SDS–PAGE and mass spectrometry. Bands A, B, C, and D were subjected N-terminal sequence analysis. (d) A diagram illustrating the products after the trypsin treatment. Two major fragments were produced from KOX1–KRAB1-90 and three fragments were yielded from the KAP-1–RBCCHIS, 22–418. The numbers represent amino acid position.

for limited proteolysis on the highly purified KOX1–KRAB(AB)1-90:KAP-1–RBCCHIS, 22–418 protein complex. Consistent digested bands were produced by these proteases for the protein complex (data not shown). The unbound KAP-1–RBCCHIS, 22–418 protein and KOX1–KRABHIS,1–90 protein were also subjected to limited proteolysis. No significant change in the digestion pattern was observed by SDS–PAGE for the unbound KAP-1–RBCCHIS, 22–418 compared to its pattern in the protein complex (data not shown). Conversely, the unbound KOX1–KRABHIS,1–90 protein was digested to a series small fragments while only partial digestion occurred in its complexed counterpart (Figure 5(c); data not shown). These data indicate that only the KOX1–KRAB protein bound to the RBCC domain of KAP-1 was protected from protease digestion. Trypsin was chosen to represent one of the limited proteolyses products from the protein complex. Mass spectrometry and N-terminal sequence analyses of selected bands of A, B, C, and D were performed after limited proteolysis (Figure 5(c)). The results show that two major fragments were produced from KOX1–KRAB(AB)1–90 (Figure 5(d)). The first fragment (KOX1, 12–90) includes the entire A box and B box. The second fragment (KOX1, 12–65) contains the complete A box and the N-terminal half of the B box. Three major fragments were produced from KAP-1–RBCCHIS, 22–418. The first fragment (KAP-1, 33–418) includes the entire RING finger, two B boxes, and coiled-coil region; the second fragment (KAP-1, 142–418) contains two B boxes and coiled-coil region; and the third fragment (KAP-1, 200–418) includes the B2 box and coiled-coil region. Guided by the program Predictprotein two protein complexes were made using the two-plasmid expression system (Table 1) including (1) KOX1–KRAB(AB)12–90 and KAP-1–RBCC50-418, and (2) KOX1–KRAB(A)12–65 and KAP-1–RBC C50-418. These two protein complexes were highly expressed and soluble. However, deletion of the RING finger and/or B1 box in RBCC resulted in unstable and insoluble proteins with or without the co-expression of KOX1–KRAB(AB)12–90. Taken together, these data suggest that the core complex may be comprised of KOX1–KRAB(AB)12–90 and the entire RBCC motif of KAP-1.

Stoichiometry of the KRAB(Ab):RBCC complex is 1:3

Although it has been demonstrated that one molecule of KRAB(AB) associates with three molecules of KAP-1–RBCC, it has not been determined whether this stoichiometry exists for KRAB(Ab). To determine the ratio of AJ18-KRAB(Ab) to KAP-1–RBCC, the two purified proteins were mixed together, and then the complex was isolated by gel filtration (Figure 6(a)). The resulting KRAB(Ab):KAP-1–RBCC protein complex was analyzed using sedimentation equilibrium. The data were best described by a model containing a complex in 1:3 molar ratio of AJ18–KRAB(Ab):KAP-1–RBCC (81%, 85%, and 94% in three protein concentrations, respectively), and small amount of aggregate (19%, 15%, and 6%, respectively) at speed of 10,000 rpm (Figure 6(b)). This result indicates that one molecule of AJ18–KRAB(Ab) directly interacts with three molecules of KAP-1–RBCC. The AJ18–KRAB(Ab) use the same mode as KOX–KRAB(AB) to bind to the KAP-1–RBCC.

Next, to determine if the secondary structure of the KRAB(Ab) is changed upon the binding of the KAP-1–RBCC, AJ18–KRAB(Ab) protein was titrated with the RBCC protein and analyzed by CD spectra. For comparison, free and bound KRAB(Ab) spectra are overlaid (Figure 7). The CD spectra show little difference between the free and bound KRAB(Ab). The data suggest that only a subtle change in secondary structure of the KRAB domain occurs upon the binding of the RBCC domain. We entertained the possibility that more sensitive and dynamic techniques may be necessary to monitor any conformation change of the KRAB domain that powerful tools for characterizing the folded state of occurs upon binding of the KAP-1–RBCC.

Potential induced structure of the KRAB domain by the RBCC domain of KAP-1

NMR and CD spectroscopies are powerful tools for characterizing the folded state of proteins in solution. Typically, a well-folded protein exhibits large ^{1}H chemical shift dispersion and a single set of resonances, for example, in the 2D ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) experiment. Figure 8 shows the HSQC spectra of His-AJ18 KRAB(Ab) (Figure 8(a) and (b)) and ZNF274-KRAB (AB) (Figure 8(d)). In general, the majority of backbone amide resonances exhibit poor dispersion, indicating both constructs have substantial regions of disordered structure and do not fold into a single globular protein conformation, consistent with CD results. This is particularly true for the ZNF274–KRAB(AB) construct, where almost all backbone resonances fall (>95%) within the center of the spectrum (7.7 ppm to 8.5 ppm). Some resonances are broadened or missing (~50 out of 75 backbone resonances can be readily detected) indicating intermediate timescale motions. This might reflect exchange between equally unfolded conformations and/or the presence of some transient local structure. It is important to note that unfolded proteins can contain nascent structure. The presence of nascent structure is consistent with the CD studies and the presence of some resonances outside the "unstructured envelope" of the HSQC spectrum. The His-AJ18-KRAB(Ab) construct, displays greater spectral heterogeneity, some of which is due to the presence of the His-tag (see below). Nevertheless, a core of sharp and broad resonances are observed for a narrow chemical shift range (~80 resonances between 7.9 ppm and 8.8 ppm in Figure 8(a) and (b)), consistent with a mostly disordered configuration. At very low contour levels, a number of very broad/weak resonances are observed, indicating the

presence of exchange between unfolded conformers characterized by some nascent structure (Figure 8(b)). Importantly, 2D nuclear Overhauser enhancement spectroscopy (NOESY) of both these samples exhibited minimal NOE cross-peaks (data not shown) providing further evidence of lack of a well defined 3D conformation. The 1D ^{1}H spectra of AJ18–KRAB(Ab), His-AJ18–KRAB(Ab) and ZNF274–KRAB(AB) (Figure 9(a), (b) and (c)) further confirm the lack of structure for the KRAB fragments. Note the absence of resonances below 0 ppm, indicating absence of ring current shifted methyl resonances (in

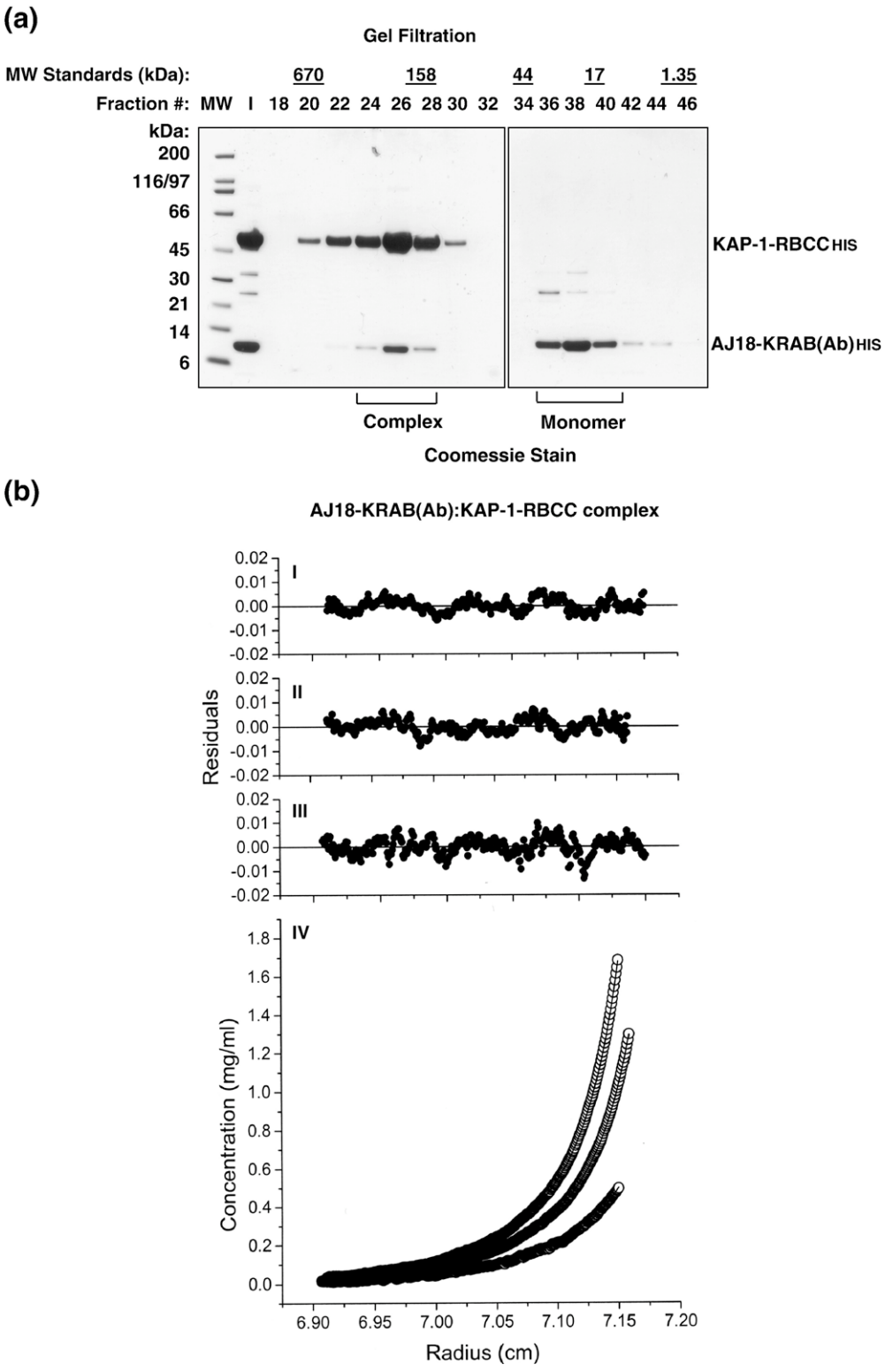


Figure 6 (*legend on next page*)

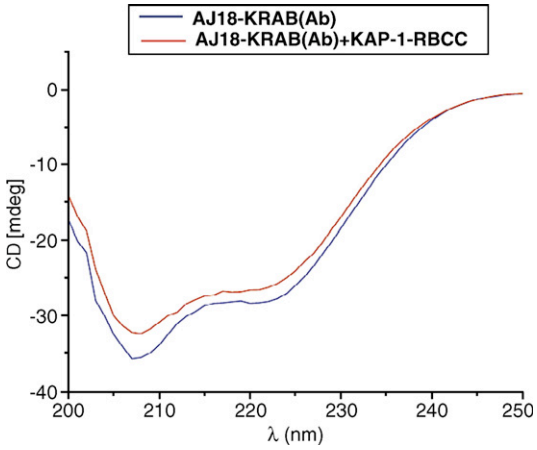


Figure 7. The CD spectra of the free and bound AJ18–KRAB(Ab) protein. CD spectra were carried out in a buffer containing 20 mM Hepes (pH 7.5), 100 mM NaCl, 1 mM TCEP at a protein concentration of 0.1 mg/ml (AJ18–KRAB(Ab)), and 0.6 mg/ml (KAP-1–RBCC), respectively. A CD spectrum for a AJ18–KRAB(Ab) protein containing 65% random coil and 35% helix is plotted (blue). A CD spectrum for a AJ18–KRAB(Ab) protein titrated with KAP-1–RBCC protein containing 65% random coil and 35% helix is plotted (red).

these constructs, there are 18/19 Ile,Val or Leu residues and 8/7 aromatic residues). Significantly, removal of the 12 residue hexa His-tag (compare Figure 9(a) and (b)), greatly reduces the ^{1}H chemical shift dispersion for the AJ18–KRAB(Ab) construct and sharpens the spectrum, indicating a reduction in any nascent structure. This is concomitant with a CD results, which show a decrease in helical structure content upon removal of the His-tag. Consistent with these studies, the proteolysis of the free KRAB constructs, clearly indicates that any such structure is transient as it does not lead to protection from proteases. Overall, the qualitative analysis of the NMR spectra, coupled with the proteolysis and CD results, indicate that the KRAB fragments, in their active form, do not exhibit properties normally associated with folded globular proteins. These data strongly suggest that the polypeptide is largely disordered, although there is evidence of limited structure that may be transiently populated and critical for activity.

To assess whether the KRAB domain folds into an ordered structure upon the formation of a complex with KAP-1, experiments were carried out using ^{15}N labeled His-AJ18–KRAB(Ab) and unlabeled KAP-1–RBCC proteins. To ensure that the KRAB(Ab) protein was indeed in a complex, the proteins were purified separately, incubated together and then complexes of the appropriate size were isolated by gel filtration (Figure 6(a)). Free and bound KRAB (Ab) fractions were then concentrated and analyzed with NMR. The ^{1}H-^{15}N HSQC spectra of the free KRAB(Ab) and the complex are shown in Figure 8(a) and (c). Only signals from the KRAB(Ab), but not from KAP-1–RBCC, were observed since unlabeled KAP-1–RBCC was used. For the KRAB(Ab) in the complex, there are fewer peaks present than evident in the free spectrum. The loss of signals is consistent with KRAB(Ab) binding to the much larger KAP-1–RBCC protein. Given that the KRAB(Ab):KAP-1–RBCC complex is 1:3 molar ratio, the complex is a 180 kDa species. At this molecular mass, molecular tumbling is slow therefore the NMR signals are too broad to be readily detected. Thus, the loss of observable peaks in the complex *versus* in the free spectra indicates that KRAB(Ab) was binding to KAP-1–RBCC. Clearly, there are some peaks remaining in the complex spectra. However, the peaks that are still present likely represent mobile NH_2 and COOH termini of KRAB(Ab) that are not making direct contacts to the KAP-1–RBCC. These data are consistent with the limited proteolysis and mass spectrometry analysis, for example these predicted three Gly residues to be visible in the NMR spectrum, which is observed. Results from total correlated spectroscopy (TOCSY) experiments were also consistent with both the NH_2 and COOH termini being flexible (data not shown). Without further analysis, which is beyond the scope of the present study, it is not possible to definitively assess the conformational state of the KRAB in complex with KAP-1–RBCC. However, the results of the proteolysis of the KRAB:KAP-1 complex, suggest that KRAB is well folded upon binding KAP-1. The presence of conformational exchange or mobile regions outside the N and C termini of KRAB would lead to a greater degree of digestion. In addition, the spectrum of KRAB in complex with KAP-1–RBCC exhibits no heterogeneity, which would be expected if some chemical exchange processes were present. Finally, some peaks were found to undergo chemical shift perturbation, which may result from a different chemical environment existing near the complex interface upon complex

Figure 6. (a) Gel filtration analysis of the AJ18–KRAB(Ab):KAP-1–RBCC protein complex. The reconstitution of AJ18–KRAB(Ab):KAP-1–RBCC protein complex *in vitro* is described in Materials and Methods. Individual proteins were purified separately and then mixed together. The mixture was subjected to gel filtration to separate unbound AJ18–KRAB (Ab) protein. The fractions were collected and the proteins were analyzed by SDS–PAGE. The peak fraction of the complex (fraction 26) eluted with a molecular mass of ~158 kDa. Significant amounts of AJ18–KRAB(Ab) was bound to the KAP-1–RBCC protein. (b) Analytical ultracentrifugation analysis of AJ18–KRAB(Ab):KAP-1–RBCC protein complex. Sedimentation equilibrium analysis was performed at 10,000 rpm and 4 °C. Panel IV shows the concentration *versus* radius data for three loading concentration of the protein complex (circles). The three data set were fitted globally with a model describing a predominant complex with 1:3 molar ratio of AJ18–KRAB(Ab):KAP-1–RBCC and small amount of aggregate. I–III shows the residuals for the data points of the fitted curves at the three protein concentrations displayed from highest concentration (top) to lowest concentration.

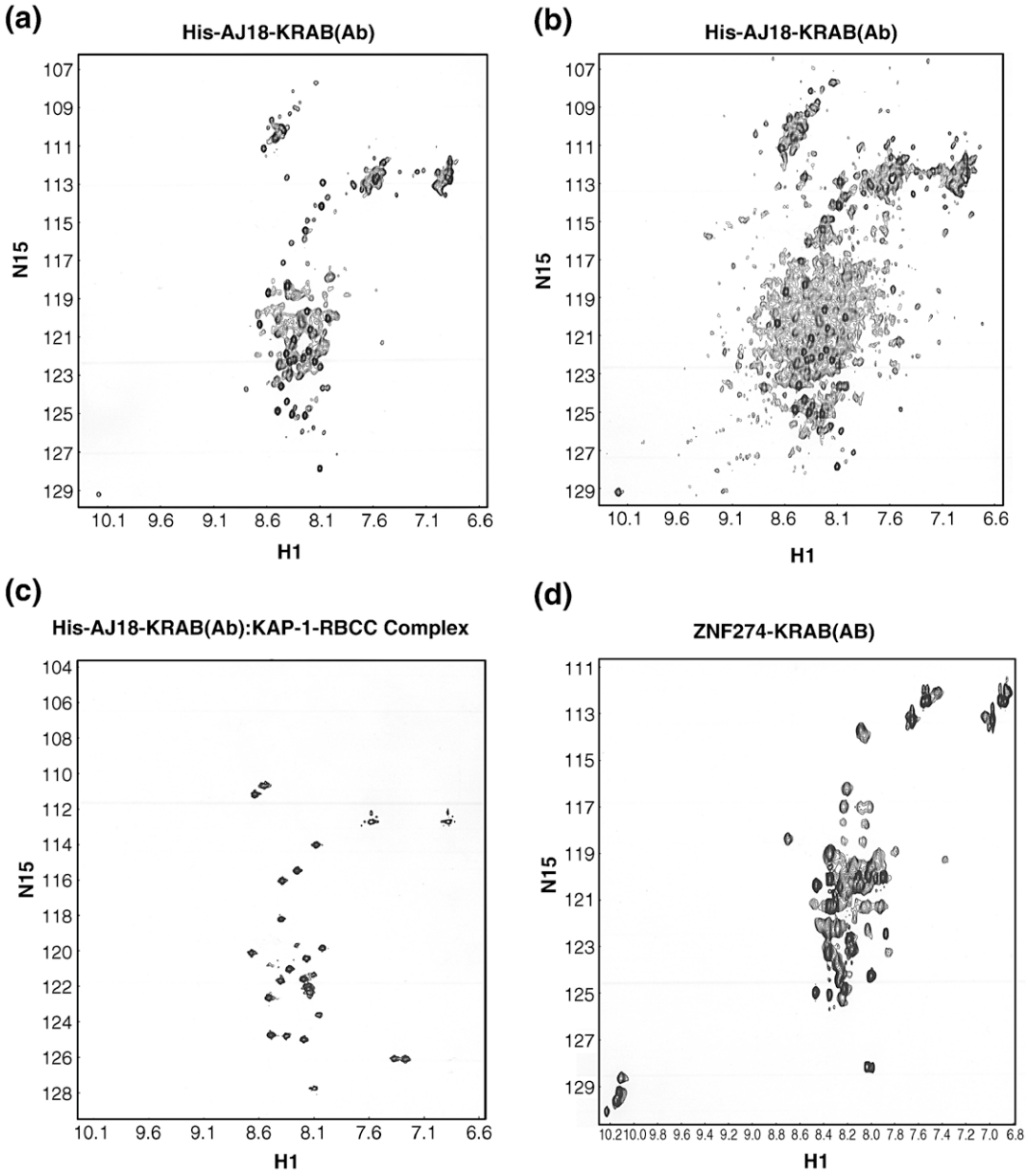


Figure 8. 2D ^{1}H-^{15}N-HSQC spectra of (a) His-AJ18–KRAB(Ab) (high contour), (b) His-AJ18–KRAB(Ab) (low contour), (c) His-AJ18–KRAB(Ab):KAP-1–RBCC complex, and (d) ZNF274–KRAB(AB).

formation, and/or due to folding. These data, taken together with the limited proteolysis data, strongly suggest that addition of the KAP-1–RBCC leads to the folding of KRAB domain.

Discussion

Here, we systematically evaluated the oligomeric states of KRAB domains from KRAB–ZNPs alone or bound in a complex with KAP-1–RBCC. We also determined their binding affinity and the stoichiometry of KRAB:KAP-1–RBCC association. We conclude the following from our data. (1) The KRAB (AB) and KRAB(Ab) domains exist as a monomer either alone or in a complex with the RBCC domain of KAP-1, while the KRAB(AB)–SCAN domain exists as a dimer. (2) The interaction between a KRAB and KAP-1–RBCC is direct and specific, and it requires only A box in KRAB(Ab) or KRAB(AC) but both A and B boxes in KRAB(AB). (3) The KRAB (Ab) has an equivalent binding capacity to KAP-1–RBCC compared to the KRAB(AB). (4) A stable complex can be formed between the KRAB and KAP-1–RBCC proteins *in vivo* and *in vitro*, and one molecule of KRAB directly interacts with three molecules of KAP-1–RBCC. (5) Limited proteolysis, mass spectrometry and N-terminal sequence analyses indicate that the core complex is formed by the entire RBCC domain of KAP-1 and the AB box of the KRAB. (6) NMR spectroscopy indicates that unbound KRAB domain is structurally disordered

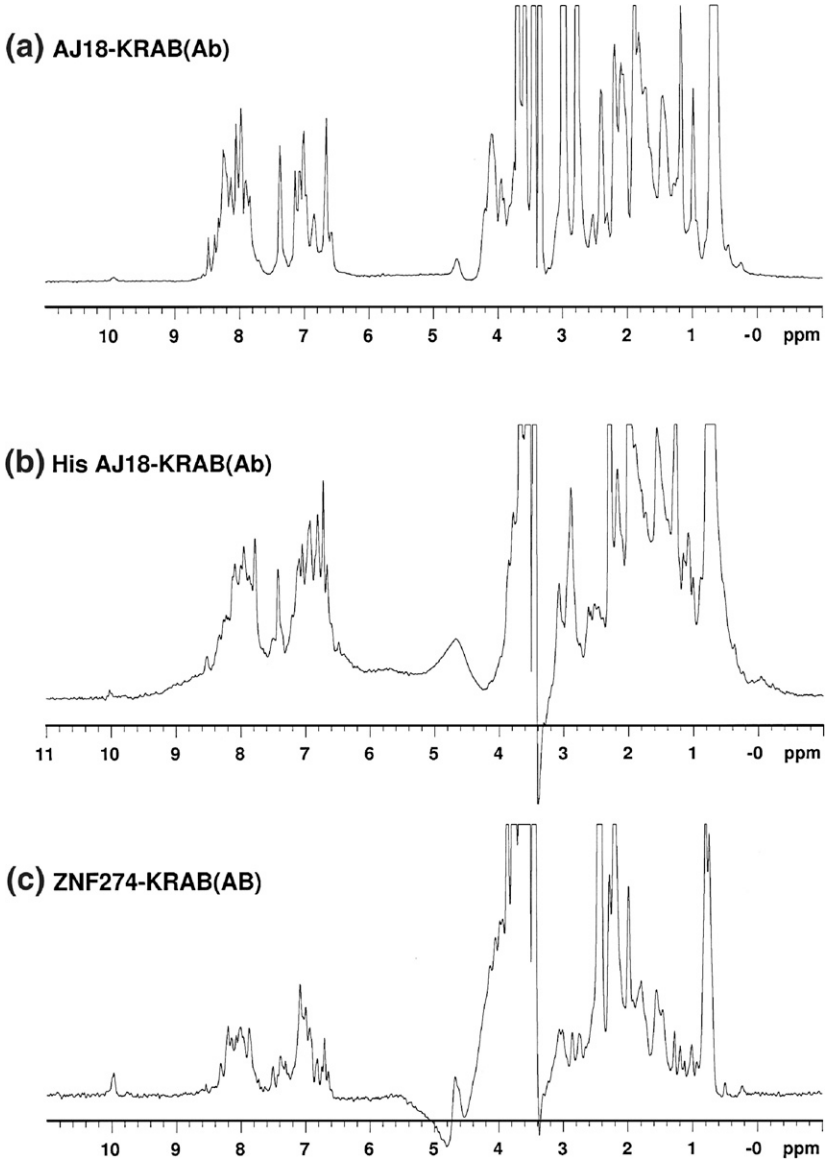


Figure 9. 1D ^{1}H spectra of (a) AJ18–KRAB(Ab), (b) His-AJ18–KRAB(Ab), and (c) ZNF274–KRAB (AB).

in solution and actively binds its target, the KAP-1-RBCC protein.

Previous studies using purified recombinant proteins indicated that the RBCC domain of KAP-1 is necessary and sufficient for interaction with KOX1–KRAB(AB), and this interaction is direct and specific.[11,12] We now demonstrate that direct interaction also exists between the KRAB(Ab), KRAB(AC), KRAB(AB)–SCAN and the KAP-1–RBCC, which is consistent with the observation from the yeast two-hybrid analysis of the interaction between different KRAB domains of the KRAB–ZFPs and the TIF1 protein family. This interaction is stable when subjected to *in vitro* biochemical manipulation as detected by GST association assay, gel filtration, and ultracentrifugation. Moreover, the binding affinity of KRAB(Ab) to RBCC is comparable to that of KRAB(AB) as determined by the competition assay. These data provide biochemical evidence in support of the proposed model that KAP-1 serves as a co-repressor for the whole KRAB–ZFP family. However, it remains unclear why the B box in the context of KRAB(AB) is necessary for binding to the RBCC, while the b or C box in the context of KRAB(Ab) or KRAB(AC) is not.

Although it has been demonstrated that the KRAB domain binding induces and/or stabilizes the trimer of the KAP-1–RBCC protein and that the KOX1–KRAB domain functions as a monomer to bind to the homotrimer of the KAP-1–RBCC protein,[11,12] little is known about its structure and the protein–protein interface of this complex. Here, we demonstrate for the first time using sedimentation equilibrium that one molecule of KRAB directly interacts with three molecules of KAP-1–RBCC in a stable highly purified protein complex. This datum is consistent with previous findings using an optical biosensor assay and quantitation analysis by densitometry *in vitro*.[11] Moreover, we confirmed by limited proteolysis, mass spectrometry and N-terminal sequence analyses that both proteins form a core complex, which is composed of the entire

RBCC domain of KAP-1 and AB box of the KRAB domain. This result is consistent with the previous observations that the RBCC domain functions as an integrated and cooperative structural unit wherein each sub-domain contributes to oligomerization and/or ligand recognition.[12]

The most important finding from this study is that the unbound KRAB domain is structurally disordered in solution and actively binds its target, the KAP-1–RBCC protein. Initially, we chose to use the KOX1–KRAB in a complex with the KAP-1–RBCC for screening crystal formation, since both proteins can be co-expressed and co-purified using a two-plasmid system to form a stable complex. However, no crystal formed from these protein complexes. We also screened other KRAB domains from the KRAB (Ab)–ZFP and KRAB(AB)–SCAN-ZFP subfamilies for crystal formation, since these proteins are highly soluble and stable in solution. Again, these KRAB domains failed to form crystals. Therefore, we applied NMR techniques to elucidate the structural and dynamic characteristics of the KRAB domain molecules. NMR has been successfully utilized to determine the functional roles and structures of the structurally disordered proteins which have basic regulatory roles in key cellular processes such as transcriptional regulation, translation and cellular signal transduction.[54–56] One of the most well characterized examples of this is the kinase-inducible transcriptional-activation domain (KID) of CREB transcriptional activator. The KID polypeptide is intrinsically disordered, as both an isolated peptide and in full-length CREB. It folds to form a pair of orthogonal helices upon binding to its target (KIX) domain in CBP transcriptional co-activator. Other examples of synergistic folding during the complex formation are the nuclear-receptor co-activator-binding domain (NCBD) of CBP and the activator for thyroid hormone and retinoid receptor domain (ACTR) of the p160 nuclear-receptor co-activator and the Mad-Sin3 repressor–co-repressor complex. Both of these protein partners undergo mutual folding transitions upon complex formation.[57] Similarly, we have found that KRAB domains are also structurally disordered and that KAP-1 association aids in folding KRAB domains.

The benefits of structurally disordered proteins have been elucidated.[38,56] One of the advantages is that unfolded protein domains undergo a significant disorder–order transition which facilitates molecular recognition. These unfolded proteins remain in an open/extended conformation when they complex with their target. This provides a disproportionately large binding surface and multiple contact points for protein–protein interaction. The KRAB domain is highly enriched in hydrophobic amino acid residues and cannot bury enough of its hydrophobic surface to form a stable globular structure in the absence of a binding partner. A molecule of the KRAB domain (M_r is 10 kDa) is much smaller than the RBCC domain of KAP-1 (M_r is 45.9 kDa), which functions as a homotrimer. Each sub-domain of the RBCC motif contributes to KRAB domain recognition. Therefore, it is very possible that the disordered structure of the KRAB domain protein facilitates the assembly of the KRAB:KAP-1–RBCC complex by forming an extensive intermolecular interface. This provides a high binding specificity and affinity (dissociation constant (K_d) = ~142 nM, a very fast association and a comparably slow dissociation rate[11]).

Whether the KAP-1–RBCC is structurally disordered has not been determined although it forms a homotrimer in unbound or bound in a complex with the KRAB. The RBCC protein is relatively unstable and tends to precipitate in solution without its binding partner, KRAB, suggesting that it might not be fully folded or that it is prone to make bodies.[58] The RBCC seems to associate with the KRAB by a mutual induced-fit mechanism: both proteins are significantly more structured in a complex than in a free state. This is consistent with our previous observation that the KRAB stabilizes the RBCC trimer.[12] The RBCC domain requires zinc for folding as detected by ICP spectrometry. One molecule of RBCC binds four zinc atoms. Eight conserved cysteine and histidine residues in the RING finger and only four conserved cysteine and histidine residues in each of the B1 and B2 boxes are used for zinc ligation. These data are well correlated with a previous finding that the mutations of conserved pairs of cysteine/cysteine and cysteine/histidine residues in RING finger and B2 box disrupt the structure of RBCC domain and concomitantly abolish the KRAB binding.[12]

Recently a protein containing only a KRAB domain (KRAB-O) was identified to interact with the SRY protein using a yeast two-hybrid screen.[59] The KRAB-O lacks a zinc finger region but contains a classical A box and B box. It has been shown to interact with the BRG region of the SRY protein *in vivo* and *in vitro*. Finding new ligands for the KRAB domain would open new doors in KRAB domain structural studies, since the SRY is much smaller than the RBCC domain. Whether the KRAB-O domain is structured has yet to be determined. If it is structurally disordered without binding to its target, then the binding of SRY might induce the structure of the KRAB domain so that this complex could be detected by NMR spectroscopy.

Taken together, all data presented here further support the current working model for KRAB–ZFP–KAP-1-mediated transcriptional repression. A KRAB–ZFP binds target gene DNA sequence-specifically through its array of C_2H_2 zinc fingers. The DNA-bound KRAB domain recruits the KAP-1 co-repressor *via* direct interaction with the RBCC domain. KAP-1 apparently self-assembles into a homotrimer, which can actively bind the KRAB domain. The HP1BD (PXVXL motif), PHD finger, and bromodomain of KAP-1 comprise the surfaces that mediate gene silencing *via* interaction with the downstream targets including the heterochromatin proteins, NURD histone deacetylase complex, SETDB1 histone methyl transferase, and other protein modification complexes (A. Ivanov, unpub-

lished data).[32–36] According to this model, we postulate that KAP-1 functions as a molecular scaffold that, in turn, coordinates the activities of large macromolecular complexes that modify chromatin structure to silence gene expression.

Materials and Methods

Preparation of plasmids

Single expression

The KRAB domain of KRAB-ZFPs were expressed using (1) IMPACT I, (2) QIAexpress, (3) GST Gene fusion, and (4) Baculovirus Express systems, respectively (Figure 1; Table 1). The pQE30 KOX1–KRAB (1–161), and pQE30 KOX1–KRAB (1–90) plasmids were described.[11,12] The pQE30 KOX1–KRAB (12–76), pQE30 SZF1–KRAB (3–110), pQE30 ZNF133–KRAB (1–119), pQE12 ZNF133–KRAB (1–119), pQE30 ZNF140–KRAB (1–119), pQE30 ZNF141–KRAB (1–107), pQE30 EEK1–KRAB (1–122), pQE30 KRK1–KRAB (16–127), and pQE30 AJ18–KRAB (4–91) plasmids were constructed *via* PCR using KRAB–ZNF cDNA as a template and primer pairs (data upon request). The PCR product was digested and cloned into the corresponding sites of the pQE vector (Qiagen). The GST–KOX1–KRAB (1–90), GST–KOX1–KRAB (DV), GST–SZF1–KRAB (3–110) plasmids were described[11,45] The GST–KOX1–KRAB (2–59), GST–AJ18–KRAB-L (4–216), GST–AJ18–KRAB (4–91), GST–AJ18–KRAB (4–58), GST–MTZ1–KRAB (2–81), GST–MTZ1–KRAB (2–50), GST–MZF13–KRAB (2–70), GST–MTZ13–KRAB (2–49), GST–ZNF202-SCAN–KRAB (37–286), GST–ZNF274–KRAB-SCAN (4–246), and GST–ZNF274–KRAB (4–77) plasmids were generated *via* PCR using KRAB–ZNF cDNA as a template and primer pairs (data upon request). The PCR product was digested and cloned into the corresponding sites of the pGEX vector (GE Healthcare).

The RBCC domain of KAP-1 was expressed using (1) QIAexpress, and (2) Baculovirus Express systems, respectively. The pQE30 KAP-1–RBCC (22–418) plasmid was described.[12] The pQE30 KAP-1–RBCC (50–418), pQE30 KAP-1–B1B2CC (136–418), and pQE30 KAP-1–B2CC (199–418) plasmids were constructed by amplifying the DNA fragments encoding desired fragment of the KAP-1 gene using the primer pairs (data upon request), and the PCR product was digested and cloned into the corresponding sites of the pQE30 vector.

Co-expression

The KRAB domain and the RBCC domain complexes were expressed using a two-plasmid express system in (1) bacteria, and (2) baculovirus, respectively. The pRM1 KOX1–KRAB (1–90), pRM1 KOX1–KRAB (12–76), pRM1 KOX1–KRAB (12–90), pRM1 KOX1–KRAB (12–65), and pRM1 AJ18–KRAB (4–91) plasmids were produced *via* PCR using KRAB–ZNF cDNA as a template and primer pairs (data upon request). The PCR product was digested and cloned into the pRM1 vector. The pQE30 KAP-1-RBCC plasmid was generated as described above.

Protein expression and purification

For single expression, the plasmids of expressing His-tagged proteins were transformed into *E. coli* SG13009 cells (Qiagen), and the plasmids of expressing GST-fusion proteins were transformed into BL21 (DE3) cells (Invitrogen). For co-expression, the plasmids of expressing non-tagged KRAB domain were first transformed into BL21 (DE3) cells, which were then chemically treated to become competent to allow the transformation of the second plasmid, pQE30 KAP-1–RBCC.

The SG13009 and BL21 (DE3) cells bearing the desired plasmid were propagated with aeration at 25 °C in either 2 X YT medium or minimal media, containing $^{15}NH_4Cl$ as the sole nitrogen source (for uniformly labeled ^{15}N protein) to an A_{600} of approximately 0.8. IPTG was added to 1 mM, and growth at 15 °C was continued (~18 h). The cells were harvested by centrifugation.

The His-tagged protein or co-expressed protein complex was purified at 4 °C under native conditions as recommended by the manufacturer (Qiagen).[12] The bacterial pellet was resuspended in the buffer A (20 mM Hepes (pH 7.5), 500 mM NaCl, 10% (v/v) glycerol) with 5 mM imidazole, and incubated with lysozyme (100 μg/ml) for 30 min in ice. Then 1 mM PMSF was added into the cell lysate, followed by sonication and centrifugation. The supernatant fraction containing soluble protein was incubated with Ni-NTA resin for 1 h. The resin was washed with buffer A with 20 mM imidazole and 1 mM PMSF, loaded into a column, and washed with buffer A in the absence of glycerol. The protein or protein complex was eluted with buffer A with 300 mM imidazole, followed by gel filtration and concentration.

GST-fusion proteins were purified at 4 °C following the manufacturere's instruction. The bacterial pellet was resuspended in buffer A with 10 mM DTT, and incubated with lysozyme (see above). Then 1 mM PMSF was added into the cell lysate, followed by sonication and centrifugation. The supernatant fraction containing soluble protein was incubated with GST resin for 2 h. The resin was washed with buffer A without PMSF, followed by thrombin digestion (Amersham) in buffer B (20 mM Hepes (pH 7.5), 100 mM NaCl, 10 mM DTT) at 4 °C (~20 h). The protein was released from GST in buffer B with 1 mM PMSF, and then purified using Q-Sepherose column in buffer B with NaCl from 100 mM to 1 M. The protein fractions were pooled and concentrated, subjected to gel filtration and concentrated to desired concentration.

Gel-filtration analysis

Gel filtration of highly purified proteins was performed using a Superdex-200 or –75 column (Amersham) equilibrated in the appropriate buffers with 1 mM TCEP, using an FPLC system as described.[12] The column was run at 4 °C and fractions were collected. Aliquots of each fraction were analyzed by SDS–PAGE.

Dynamic light scattering

All measurements were carried out on a DynaPro-801 molecular sizing instrument at 4 °C. Protein samples (5 mg/ml) were run and pre-filtered through a 0.02 μm membrane filter (Whatman, Anodisc 13) prior to analysis.

Circular dichroism spectroscopy

Spectra were collected on a Jasco 810 spectropolarimeter at 25° using a 1 mm path-length cuvette (Hellma).

Spectra were collected over the wavelength range 200 nm–250 nm using a resolution of 0.5 nm and bandwidth of 1 nm at room temperature. Each scan was taken at 10 nm min^{-1} with a response time of 8 s. Spectra were baseline corrected for buffer contributions and mean residue ellipticity was calculated using the program JFIT‡. Concentrations of AJ18–KRAB(Ab), 274-KRAB–SCAN, and KAP-1–RBCCHIS were 0.1 mg/ml, 0.2 mg/ml, and 0.6 mg/ml, respectively.

Analytical ultracentrifugation

Sedimentation equilibrium experiments were performed as described previously.[12] The experiments were performed at 4 °C using a minimum of two speeds. The fringe displacement data was collected every 4–6 hours until equilibrium was reached, as determined by comparing successive scans using the MATCH v.7 program[1]. Data was edited using REEDIT v.9 program[1] and analyzed using the NONLIN v.3 program[1].[46] The program SEDNTERP was used to calculate the molecular weight and the partial specific volume of the proteins based on the amino acid sequence. This program was also used to calculate the solvent density. Three data sets from different initial loading concentrations were fitted simultaneously. Examination of the residuals and minimization of the variance determined the goodness of fit.

[1]MATCH, REEDIT, and NONLIN are available from the Analytical Ultracentrifugation Facility at the University of Connecticut *via* the FTP site. The program SEDNTERP was written by T. Laue, J. Hayes, and J. Philo and is available on the RASMB web site.

GST association assays

The GST association assays were performed as described.[11]

Competition assays

A total of 5 μg of freshly prepared GST-KOX1–KRAB (AB) fusion protein immobilized on GST resin was mixed with 1 μg, 5 μg, and 10 μg of KOX1–KRAB(AB), AJ18–KRAB(Ab), and ZNF274–KRAB(AB)–SCAN proteins in 200 μl of BB500 buffer (20 mM Tris (pH 7.9), 500 mM NaCl, 0.2 mM EDTA, 10% glycerol, 0.2% (v/v) Nonidet P-40, 1 mM PMSF) with 1 mg of bovine serum albumin. The 5 μg of KAP-1–RBCCHIS protein were then added into reaction and incubated for 1 h at room temperature. The protein complexes were washed four times with BB750 (same as BB500 with 750 mM NaCl), and the bound proteins were eluted in 5 X Laemmli buffer, resolved by SDS–PAGE, and visualized with Coomassie blue stain.

ICP spectrometry

ICP spectrometry was conducted at Galbraith Laboratories, Inc. (Knoxville, TN). Two hundred μl of a 200 μM AJ18–KRAB(Ab)HIS sample (as negative control for non-specific HIS tag binding to zinc), 150 μl of 330 μM, and 150 μl of 33 μM AJ18–KRAB(Ab)HIS: KAP-1–RBCCHIS complex (PBS, 500 mM NaCl, 1 mM TCEP (pH 7.0)) were prepared for ICP analysis using a wet ash digestion procedure. Measurements were made on a Perkin-Elmer P2000 using a primary wavelength of 213.856.

Limited proteolysis, mass spectrometry, and N-terminal sequence analyses

The protein complex was treated with trypsin protease (sequencing grade modified trypsin; Promega) in buffer C (20 mM Hepes (pH 7.5), 100 mM NaCl, 1 mM β-mecaptoethanol) in ice for 60 min. The concentration of trypsin in the reaction was 0.04 μg/μl and 0.008 μg/μl, respectively. The reaction was terminated with 5 mM DFP (trypsin protease inhibitor) in ice for 48 h. Aliquots of each reaction were analyzed by SDS–PAGE, mass spectrometry, and N-terminal sequence analyses.

Reconstitution of the AJ18-KRAB(Ab):KAP-1–RBCC complexes *in vitro*

To reconstitute the complex, the AJ18–KRAB(Ab) and the KAP-1–RBCCHIS proteins were purified first under native conditions as described above. Both proteins were then mixed at 4 °C for 1 h. The protein complex was analyzed by gel filtration and SDS–PAGE.

NMR spectroscopy

NMR spectra were acquired at 25 °C on a Varian 600 MHz spectrometer. 1D ^{1}H and 2D NOESY spectra were acquired using the watergate sequence for solvent suppresion. Two-dimensional ^{1}H-^{15}N HSQC spectra were typically acquired under the following conditions: protein or protein complex concentrations were 1 mM in buffer PBS, 500 mM NaCl, 1 mM TCEP and 10% (v/v) $^{2}H_2O$/90%H_2O. Spectra were acquired with either 32 scans or 128 scans and processed using NMRPipe.[60]

Acknowledgements

F.J.R. is supported, in part, by National Institutes of Health grants CA 095561, CA 092088, and DAMD17-02-1-0631. We thank L. Hellman for the MZF13 and MTZ1 plasmids, A. Jheon for the AJ18 plasmid, C. Schumacher for ZNF202 plasmid, W. J. Fredericks, D. C. Schultz, M. S. Lechner, and A. Ivanov for helpful discussions, I. Topisirovic for helping with the CD experiments. We thank the Wistar Institute DNA Core Facility for DNA sequencing, Protein Core Facility for baculovirus protein expression, Protein Microchemistry/MS Facility for MASS spectrometry and N-terminal sequence analyses, and Commonwealth Universal Research Enhancement Program, Pennsylvania Department of Health.

‡ ftp://jgiqc.llnl.gov/cdfit/

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2007.03.047

References

1. Barwell, J. A., Bochkarev, A., Pfuetzner, R. A., Tong, H., Yang, D. S., Frappier, L. & Edwards, A. M. (1995). Overexpression, purification, and crystallization of the DNA binding and dimerization domains of the Epstein-Barr virus nuclear antigen 1. *J. Biol. Chem.* **270**, 20556–20559.
2. Huynh, K. D. & Bardwell, V. J. (1998). The BCL-6 POZ domain and other POZ domains interact with the co-repressors N-CoR and SMRT. *Oncogene*, **17**, 2473–2484.
3. Li, X., Peng, H., Schultz, D. C., Lopez-Guisa, J. M., Rauscher, F. J., 3rd & Marmorstein, R. (1999). Structure-function studies of the BTB/POZ transcriptional repression domain from the promyelocytic leukemia zinc finger oncoprotein. *Cancer Res.* **59**, 5275–5282.
4. Ahmad, K. F., Engel, C. K. & Prive, G. G. (1998). Crystal structure of the BTB domain from PLZF. *Proc. Natl Acad. Sci. USA*, **95**, 12123–12128.
5. Williams, A. J., Blacklow, S. C. & Collins, T. (1999). The zinc finger-associated SCAN box is a conserved oligomerization domain. *Mol. Cell Biol.* **19**, 8526–8535.
6. Schumacher, C., Wang, H., Honer, C., Ding, W., Koehn, J., Lawrence, Q. *et al.* (2000). The SCAN domain mediates selective oligomerization. *J. Biol. Chem.* **275**, 17173–17179.
7. Stone, J. R., Maki, J. L., Blacklow, S. C. & Collins, T. (2002). The SCAN domain of ZNF174 is a dimer. *J. Biol. Chem.* **277**, 5448–5452.
8. Ivanov, D., Stone, J. R., Maki, J. L., Collins, T. & Wagner, G. (2005). Mammalian SCAN domain dimer is a domain-swapped homolog of the HIV capsid C-terminal domain. *Mol. Cell*, **17**, 137–143.
9. Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X. P., Neilson, E. G. & Rauscher, F. J., 3rd (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev.* **10**, 2067–2078.
10. Margolin, J. F., Friedman, J. R., Meyer, W. K., Vissing, H., Theisen, H. J. & Rauscher, F. J., 3rd (1994). Kruppel-associated boxes are potent transcriptional repression domains. *Proc. Natl Acad. Sci. USA*, **91**, 4509–4513.
11. Peng, H., Begg, G. E., Harper, S. L., Friedman, J. R., Speicher, D. W. & Rauscher, F. J., 3rd (2000). Biochemical analysis of the Kruppel-associated box (KRAB) transcriptional repression domain. *J. Biol. Chem.* **275**, 18000–18010.
12. Peng, H., Begg, G. E., Schultz, D. C., Friedman, J. R., Jensen, D. E., Speicher, D. W. & Rauscher, F. J., 3rd (2000). Reconstitution of the KRAB-KAP-1 repressor complex: a model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein-protein interactions. *J. Mol. Biol.* **295**, 1139–1162.
13. Huntley, S., Baggott, D. M., Hamilton, A. T., Tran-Gyamfi, M., Yang, S., Kim, J. *et al.* (2006). A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. *Genome Res.* **16**, 669–677.
14. Mark, C., Abrink, M. & Hellman, L. (1999). Comparative analysis of KRAB zinc finger proteins in rodents and man: evidence for several evolutionarily distinct subfamilies of KRAB zinc finger genes. *DNA Cell Biol.* **18**, 381–396.
15. Vissing, H., Meyer, W. K., Aagaard, L., Tommerup, N. & Thiesen, H. J. (1995). Repression of transcriptional activity by heterologous KRAB domains present in zinc finger proteins. *FEBS Letters*, **369**, 153–157.
16. Bellefroid, E. J., Marine, J. C., Ried, T., Lecocq, P. J., Riviere, M., Amemiya, C. *et al.* (1993). Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells. *EMBO J.* **12**, 1363–1374.
17. Liu, C., Levenstein, M., Chen, J., Tsifrina, E., Yonescu, R., Griffin, C. *et al.* (1999). SZF1: a novel KRAB-zinc finger gene expressed in CD34+ stem/progenitor cells. *Exp. Hematol.* **27**, 313–325.
18. Wagner, S., Hess, M. A., Ormonde-Hanson, P., Malandro, J., Hu, H., Chen, M. *et al.* (2000). A broad role for the zinc finger protein ZNF202 in human lipid metabolism. *J. Biol. Chem.* **275**, 15685–15690.
19. Jheon, A. H., Ganss, B., Cheifetz, S. & Sodek, J. (2001). Characterization of a novel KRAB/C2H2 zinc finger transcription factor involved in bone development. *J. Biol. Chem.* **276**, 18282–18289.
20. Krebs, C. J., Larkins, L. K., Price, R., Tullis, K. M., Miller, R. D. & Robins, D. M. (2003). Regulator of sex-limitation (Rsl) encodes a pair of KRAB zinc-finger genes that control sexually dimorphic liver gene expression. *Genes Dev.* **17**, 2664–2674.
21. Crew, A. J., Clark, J., Fisher, C., Gill, S., Grimer, R., Chand, A. *et al.* (1995). Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with homology to the Kruppel-associated box in human synovial sarcoma. *EMBO J.* **14**, 2333–2340.
22. Tommerup, N., Aagaard, L., Lund, C. L., Boel, E., Baxendale, S., Bates, G. P. *et al.* (1993). A zinc-finger gene ZNF141 mapping at 4p16.3/D4S90 is a candidate gene for the Wolf-Hirschhorn (4p-) syndrome. *Hum. Mol. Genet.* **2**, 1571–1575.
23. Shoichet, S. A., Hoffmann, K., Menzel, C., Trautmann, U., Moser, B., Hoeltzenbein, M. *et al.* (2003). Mutations in the ZNF41 gene are associated with cognitive deficits: identification of a new candidate for X-linked mental retardation. *Am. J. Hum. Genet.* **73**, 1341–1354.
24. Kleefstra, T., Yntema, H. G., Oudakker, A. R., Banning, M. J., Kalscheuer, V. M., Chelly, J. *et al.* (2004). Zinc finger 81 (ZNF81) mutations associated with X-linked mental retardation. *J. Med. Genet.* **41**, 394–399.
25. Le Douarin, B., Zechel, C., Garnier, J. M., Lutz, Y., Tora, L., Pierrat, P. *et al.* (1995). The N-terminal part of TIF1, a putative mediator of the ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-raf in the oncogenic protein T18. *EMBO J.* **14**, 2020–2033.
26. Venturini, L., You, J., Stadler, M., Galien, R., Lallemand, V., Koken, M. H. *et al.* (1999). TIF1gamma, a novel member of the transcriptional intermediary factor 1 family. *Oncogene*, **18**, 1209–1217.
27. Khetchoumian, K., Teletin, M., Mark, M., Lerouge, T., Cervino, M., Oulad-Abdelghani, M. *et al.* (2004). TIF1delta, a novel HP1-interacting member of the transcriptional intermediary factor 1 (TIF1) family

expressed by elongating spermatids. *J. Biol. Chem.* **279**, 48329–48341.
28. Beckstead, R., Ortiz, J. A., Sanchez, C., Prokopenko, S. N., Chambon, P., Losson, R. & Bellen, H. J. (2001). Bonus, a Drosophila homolog of TIF1 proteins, interacts with nuclear receptors and can inhibit betaFTZ-F1-dependent transcription. *Mol. Cell*, **7**, 753–765.
29. Dupont, S., Zacchigna, L., Cordenonsi, M., Soligo, S., Adorno, M., Rugge, M. & Piccolo, S. (2005). Germ-layer specification and control of cell growth by Ectodermin, a Smad4 ubiquitin ligase. *Cell*, **121**, 87–99.
30. Saurin, A. J., Borden, K. L., Boddy, M. N. & Freemont, P. S. (1996). Does this have a familiar RING? *Trends Biochem. Sci.* **21**, 208–214.
31. Borden, K. L. (1998). RING fingers and B-boxes: zinc-binding protein-protein interaction domains. *Biochem. Cell Biol.* **76**, 351–358.
32. Ryan, R. F., Schultz, D. C., Ayyanathan, K., Singh, P. B., Friedman, J. R., Fredericks, W. J. & Rauscher, F. J., 3rd (1999). KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol. Cell Biol.* **19**, 4366–4378.
33. Lechner, M. S., Begg, G. E., Speicher, D. W. & Rauscher, F. J., 3rd (2000). Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: direct chromoshadow domain-KAP-1 corepressor interaction is essential. *Mol. Cell Biol.* **20**, 6449–6465.
34. Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* **16**, 919–932.
35. Schultz, D. C., Friedman, J. R. & Rauscher, F. J., 3rd (2001). Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. *Genes Dev.* **15**, 428–443.
36. Cammas, F., Herzog, M., Lerouge, T., Chambon, P. & Losson, R. (2004). Association of the transcriptional corepressor TIF1beta with heterochromatin protein 1 (HP1): an essential role for progression through differentiation. *Genes Dev.* **18**, 2147–2160.
37. Mannini, R., Rivieccio, V., D'Auria, S., Tanfani, F., Ausili, A., Facchiano, A. *et al.* (2006). Structure/function of KRAB repression domains: structural properties of KRAB modules inferred from hydrodynamic, circular dichroism, and FTIR spectroscopic analyses. *Proteins: Struct. Funct. Genet.* **62**, 604–616.
38. Wright, P. E. & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321–331.
39. Radhakrishnan, I., Perez-Alvarado, G. C., Dyson, H. J. & Wright, P. E. (1998). Conformational preferences in the Ser133-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB. *FEBS Letters*, **430**, 317–322.
40. Radhakrishnan, I., Perez-Alvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R. & Wright, P. E. (1997). Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell*, **91**, 741–752.
41. Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J. & Pavletich, N. P. (1996). Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science*, **274**, 948–953.
42. Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J. & Verdine, G. L. (1997). Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science*, **277**, 1310–1313.
43. Rosati, M., Marino, M., Franze, A., Tramontano, A. & Grimaldi, G. (1991). Members of the zinc finger protein gene family sharing a conserved N-terminal module. *Nucl. Acids Res.* **19**, 5661–5667.
44. Abrink, M., Ortiz, J. A., Mark, C., Sanchez, C., Looman, C., Hellman, L. *et al.* (2001). Conserved interaction between distinct Kruppel-associated box domains and the transcriptional intermediary factor 1 beta. *Proc. Natl Acad. Sci. USA*, **98**, 1422–1426.
45. Peng, H., Zheng, L., Lee, W. H., Rux, J. J. & Rauscher, F. J., 3rd (2002). A common DNA-binding site for SZF1 and the BRCA1-associated zinc finger protein, ZBRK1. *Cancer Res.* **62**, 3773–3781.
46. Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981). Analysis of data from the analytical ultracentrifuge by non-linear least-squares techniques. *Biophys. J.* **36**, 575–588.
47. Agata, Y., Matsuda, E. & Shimizu, A. (1999). Two novel Kruppel-associated box-containing zinc-finger proteins, KRAZ1 and KRAZ2, repress transcription through functional interaction with the corepressor KAP-1 (TIF1beta/KRIP-1). *J. Biol. Chem.* **274**, 16412–16422.
48. Barlow, P. N., Luisi, B., Milner, A., Elliott, M. & Everett, R. (1994). Structure of the C3HC4 domain by ^{1}H-nuclear magnetic resonance spectroscopy. A new structural class of zinc-finger. *J. Mol. Biol.* **237**, 201–211.
49. Bellon, S. F., Rodgers, K. K., Schatz, D. G., Coleman, J. E. & Steitz, T. A. (1997). Crystal structure of the RAG1 dimerization domain reveals multiple zinc-binding motifs including a novel zinc binuclear cluster. *Nature Struct. Biol.* **4**, 586–591.
50. Borden, K. L., Boddy, M. N., Lally, J., O'Reilly, N. J., Martin, S., Howe, K. *et al.* (1995). The solution structure of the RING finger domain from the acute promyelocytic leukaemia proto-oncoprotein PML. *EMBO J.* **14**, 1532–1541.
51. Reddy, B. A. & Etkin, L. D. (1991). A unique bipartite cysteine-histidine motif defines a subfamily of potential zinc-finger proteins. *Nucl. Acids Res.* **19**, 6330.
52. Wolf, E., Kim, P. S. & Berger, B. (1997). MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci.* **6**, 1179–1189.
53. Borden, K. L., Lally, J. M., Martin, S. R., O'Reilly, N. J., Etkin, L. D. & Freemont, P. S. (1995). Novel topology of a zinc-binding domain from a protein involved in regulating early *Xenopus* development. *EMBO J.* **14**, 5947–5956.
54. Dyson, H. J. & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nature Rev. Mol. Cell Biol.* **6**, 197–208.
55. Dyson, H. J. & Wright, P. E. (2005). Elucidation of the protein folding landscape by NMR. *Methods Enzymol.* **394**, 299–321.
56. Tompa, P. (2002). Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27**, 527–533.

57. Brubaker, K., Cowley, S. M., Huang, K., Loo, L., Yochum, G. S., Ayer, D. E. *et al.* (2000). Solution structure of the interacting domains of the Mad-Sin3 complex: implications for recruitment of a chromatin-modifying complex. *Cell*, **103**, 655–665.
58. Kentsis, A., Gordon, R. E. & Borden, K. L. (2002). Self-assembly properties of a model RING domain. *Proc. Natl Acad. Sci. USA*, **99**, 667–672.
59. Oh, H. J., Li, Y. & Lau, Y. F. (2005). Sry associates with the heterochromatin protein 1 complex by interacting with a KRAB domain protein. *Biol. Reprod.* **72**, 407–415.
60. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.

Edited by M. Yaniv

(Received 27 November 2006; received in revised form 5 February 2007; accepted 19 March 2007)
Available online 24 March 2007

Ajuba LIM proteins are Snail corepressors required for neural crest development in Xenopus

Ellen M. Langer[1], Yungfeng Feng[1], Zhaoyuan Hou[3], Frank J. Rauscher, III[3], Kristen L. Kroll[2], and Gregory D. Longmore*[1]

Departments of Medicine and Cell Biology[1] and Molecular Biology and Pharmacology[2], Washington University, Saint Louis, MO, and The Wistar Institute, Philadelphia, PA[3]

Running Title: Ajuba LIM proteins are Snail corepressors

Keywords: LIM proteins, Snail, Neural crest development, Epithelial Mesenchymal Transitions, E-cadherin, Xenopus development

Text: 62,569 characters with spaces

Abstract: 200 words

* Corresponding Author

Division of Hematology

Washington University School of Medicine, Campus Box 8125

660 South Euclid Avenue

St. Louis MO 63110

Phone: 314-362-8834

FAX: 314-362-8826

E-mail: glongmor@im.wustl.edu

Abstract

Snail family transcriptional repressors play an essential role in epithelial mesenchymal transitions (EMT) during both physiological and pathological processes. A conserved SNAG repression domain in vertebrate Snail family proteins is necessary for assembly of a repressor complex, but the mechanism by which this complex assembles is unknown. Here, we identify the Ajuba family of LIM proteins as corepressors of the Snail family through interaction with the SNAG domain. Ajuba LIM proteins interact with Snail proteins in the nucleus, contribute to Snail-dependent repression of E-cadherin, and are present on endogenous E-cadherin promoters in the presence of Snail. Using Xenopus neural crest as a model of *in vivo* Snail- or Slug-induced EMT, we demonstrate that Ajuba LIM proteins contribute to neural crest development and that blocking interaction of Snail proteins with Ajuba LIM proteins causes a loss or reduction in neural crest development. Together, these data indicate that Ajuba LIM proteins are important Snail corepressors required for *in vivo* Snail function. Because Ajuba LIM proteins contribute to junction assembly or stability as components of adherens junctions in epithelia, the functional interaction between Ajuba LIM proteins and Snail proteins in the nucleus suggests that Ajuba LIM proteins are important regulators of epithelia dynamics.

Introduction

Developmental processes such as gastrulation and neural crest cell delamination require epithelial cells to undergo a mesenchymal transition in order to invade and migrate (Thiery and Sleeman 2006). This process of an epithelial to mesenchymal transition (EMT) occurs not only during development, but also during wound repair (Savagner et al. 2005), chronic inflammation and fibrosis (Kalluri and Neilson 2003), and tumor progression from localized epithelial adenomas to metastatic carcinomas (Thiery 2002). While many environmental signals can induce EMT, they all converge to activate transcription factors that effect an EMT "gene program." This program requires transcriptional down-regulation of epithelial genes, particularly cell-cell adhesive receptors such as E-cadherin, and up-regulation of mesenchymal genes such as fibronectin. Together, these changes allow epithelial cells to lose adherence to neighboring cells and begin to migrate and invade. A fundamental challenge is to understand whether and how cell surface adhesive events communicate with nuclear processes to coordinate this dynamic transition.

Many transcription factors contribute to EMT through direct repression of epithelial genes such as the adhesive receptors E-cadherin, claudins, and occludins. These include dEF1/ZEB1 (Guaita et al. 2002), SIP1/ZEB2 (Comijn et al. 2001), E47 (Perez-Moreno et al. 2001), Twist (Yang et al. 2004), Snail (Batlle et al. 2000; Cano et al. 2000; Ikenouchi et al. 2003) and Slug (Hajra et al. 2002). Snail and Slug belong to the Snail family of zinc-finger transcriptional repressors and have been well-established as central regulators of EMT in gastrulation (Carver et al. 2001; Hemavathy et al. 2004), neural crest induction and delamination (Nieto et al. 1994; LaBonne and Bronner-Fraser 2000; del Barrio and Nieto 2002; Aybar et al. 2003), skin wound repair (Savagner et al. 2005), mesothelial fibrosis (Barrallo-Gimeno and

Nieto 2005), and tumor metastasis (Blanco et al. 2002; Barrallo-Gimeno and Nieto 2005; Moody et al. 2005). Specifically, the repressor function of Snail is necessary for Drosophila gastrulation (Carver et al. 2001; Hemavathy et al. 2004) and repression by both Snail and Slug is necessary for Xenopus neural crest induction and delamination (LaBonne and Bronner-Fraser 2000; Aybar et al. 2003). On a cellular level, repressor activities of Snail and Slug influence cell proliferation and survival (Kajita et al. 2004; Vega et al. 2004), cell adhesion, and cell migration (reviewed in (Barrallo-Gimeno and Nieto 2005). Snail and Slug affect these cellular processes both in the presence and absence of a concomitant EMT (Barrallo-Gimeno and Nieto 2005).

While Snail and Slug play integral roles in development and disease through gene repression, the precise mechanism of Snail-dependent repression remains unclear. It has been shown that repression by both Snail and Slug is sensitive to TSA, a histone deacetylase (HDAC) inhibitor (Hemavathy et al. 2000; Peinado et al. 2004). Indeed, Snail recruits an HDAC repressor complex (Peinado et al. 2004), but precisely how it assembles this complex at select promoters is unknown. In Drosophila, Snail associates with a corepressor, CtBP, which is necessary to mediate its repressor activity (Nibu et al. 1998). Vertebrate Snail family members, however, do not contain a conserved CtBP binding domain. Instead, they mediate repression through a 7-32 amino acid N-terminal SNAG (Snail-Gfi-1) domain that is necessary and sufficient for repression (Grimes et al. 1996; Nakayama et al. 1998; Hemavathy et al. 2000). In addition, the SNAG domain is necessary for recruitment of HDAC proteins and assembly of a repressor complex (Peinado et al. 2004). Together, these data strongly support the presence of cellular co-repressor(s) that interact with the SNAG domain of vertebrate Snail, analogous to CtBP and Drosophila Snail.

In a screen to identify SNAG domain interacting co-repressors, we identified the Ajuba LIM protein family. Ajuba LIM proteins (Ajuba, LIMD1, WTIP) are closely related to Zyxin LIM proteins (Zyxin, LPP, Trip6). The Ajuba/Zyxin family is characterized by three homologous C-terminal LIM domains (LIM region) and a unique N-terminal region (preLIM region). These proteins localize to cell-cell or cell-matrix adhesion sites in epithelial and fibroblast cells, respectively, and influence cell adhesive complex formation and function (Marie et al. 2003; Srichai et al. 2004; Pratt et al. 2005; Hansen and Beckerle 2006; Hoffman et al. 2006). In addition, they shuttle to and from the nucleus, suggesting they have the potential to coordinate cell surface adhesive events with nuclear responses (Nix and Beckerle 1997; Kanungo et al. 2000; Sharp et al. 2004; Srichai et al. 2004).

We show that Ajuba LIM proteins, but not related Zyxin LIM proteins, specifically interact with Snail proteins in the nucleus and function as Snail corepressors to downregulate E-cadherin transcription. Ajuba LIM proteins are recruited to the endogenous E-cadherin promoter in a Snail-dependent manner. *In vivo*, Ajuba and LIMD1 cooperate with Snail and Slug during Xenopus neural crest development and the expression pattern of endogenous LIMD1 parallels that of Snail and Slug during Xenopus development. Thus, in addition to regulating epithelial cell-cell adhesion, Ajuba LIM proteins contribute to epithelial-mesenchymal transitions as Snail co-repressors during neural crest development.

Results

Ajuba LIM proteins interact with the Snail family of proteins in the nucleus

The SNAG domain of vertebrate Snail proteins is both necessary and sufficient for transcriptional repressor activity (Nakayama et al. 1998; Hemavathy et al. 2000), but precisely how the SNAG domain mediates assembly of a repressor complex remains unclear. To identify potential Snail family co-repressors, we performed a yeast two-hybrid protein-protein interactive screen using the Gfi-1 SNAG domain fused to the LexA DNA binding domain as bait. From this screen, two of the three Ajuba family LIM proteins (Ajuba and LIMD1) were identified as potential interactors. Controls in yeast cells and mammalian cell lines demonstrated that both LIM proteins specifically interacted with a functional SNAG domain. The biochemical requirements of the interaction between Ajuba and the SNAG domain and the effect on repression as modeled by a synthetic SNAG domain-containing repressor have been analyzed and submitted as a separate manuscript (K. Ayyanathan et al. submitted).

To determine whether Ajuba LIM proteins associate with full-length Snail proteins in vertebrate cells, HEK293 cells were cotransfected with epitope-tagged LIM and Snail proteins. Snail proteins were immunoprecipitated from total cell lysates and bound products Western blotted for the presence of LIM protein. Ajuba, LIMD1, and WTIP were all found to interact with all Snail family members tested (Snail1, Slug (Snail2), and Scratch) (Fig. 1A, top and 1B). However, closely related LIM proteins, Zyxin and LPP, did not interact with Snail or its family members (Fig. 1A and 1B), indicating that the interaction with Snail proteins was specific to the Ajuba subfamily.

Next, we determined the region of Ajuba necessary for binding to Snail by co-transfecting HEK293 cells with myc-tagged full-length Ajuba, N-terminal preLIM region of

Ajuba, or C-terminal LIM region of Ajuba (Fig. 1C) and Flag-tagged Snail. Cells were lysed, Snail immunoprecipitated, and bound products Western blotted for the presence of Ajuba isoforms. The C-terminal LIM region of Ajuba, but not the preLIM region, interacted with Snail (Fig. 1C). Conversely, to define whether region(s) of Snail other than the SNAG domain also contribute to its interaction with Ajuba, we cotransfected HEK293 cells with HA-tagged full-length Snail or Snail lacking the 7 amino acid SNAG domain (ΔSNAG) with myc-tagged Ajuba (Fig. 1D). While loss of the SNAG domain decreased the interaction by 80%, Snail ΔSNAG still bound to Ajuba, suggesting that other regions of Snail could contribute to the interaction (Fig. 1D). Further determination of the Snail region(s) that contribute to the association with Ajuba was complicated by instability of N-terminal and C-terminal deletions, as has been previously observed (Peinado et al. 2004). Nonetheless, this analysis demonstrated that the major domain in Snail that mediated its interaction with Ajuba was the SNAG domain.

To confirm that the interaction of Ajuba and Snail occurred in cells containing endogenous levels of each protein we used two epithelial cell lines that express endogenous Snail: HaCaT human keratinocytes and MDA-231 human breast cancer cells. Snail is a labile protein due to proteosomal degradation following phosphorylation by GSK3β (Zhou et al. 2004). Therefore, we treated cells with MG132, a proteasome inhibitor, and LiCl, a GSK-3 inhibitor, for five hours to stabilize endogenous Snail (Zhou et al. 2004). Lysates were immunoprecipitated with a Snail antibody and bound products Western blotted for the presence of Ajuba and LIMD1. Ajuba and LIMD1 co-immunoprecipitated with Snail in both cell types (Fig. 1E), indicating that the interaction occurred in cells with endogenous levels of each protein.

After confirming that Ajuba and Snail interact, we sought to determine whether these proteins co-localize in cells. Both Ajuba and Snail shuttle to and from the nucleus (Kanungo et

al. 2000; Dominguez et al. 2003; Zhou et al. 2004). When GFP-Snail or RFP-Ajuba was expressed alone in MCF-7 human breast epithelial cells, GFP-Snail was predominantly nuclear (95% of cells) while RFP-Ajuba was predominantly cytosolic (85-90% of cells). To determine if co-expression of Snail affected the localization of Ajuba, we cotransfected MCF-7 cells with RFP-Ajuba in combination with GFP or GFP-Snail. While the percent of cells with nuclear GFP-Snail remained the same in the presence or absence of Ajuba, Snail expression caused a greater than three-fold increase in the percent of cells with nuclear RFP-Ajuba. Co-transfection of RFP-Ajuba and GFP resulted in 14% of cells with nuclear Ajuba (similar to RFP-Ajuba alone), while cotransfection of RFP-Ajuba and GFP-Snail resulted in 53% of cells with nuclear Ajuba (Fig. 2A and 2B). Snail 8 Ser-Ala (Sn8SA) is a mutant form of Snail that cannot be phosphorylated by nuclear GSK3β and, thus, is stabilized and constitutively localized to the nucleus (Zhou et al. 2004). When GFP-Sn8SA and RFP-Ajuba were co-expressed, an even greater percentage of cells (76%) contained nuclear Ajuba (Fig. 2B). Since Sn8SA does not exit the nucleus, this result suggested that Snail was capable of trapping Ajuba in the nucleus.

In another approach to assess the interaction of Ajuba and Snail in the nucleus, HaCaT cells stably expressing myc-Ajuba were fractionated into nuclear and cytosolic extracts. Snail was immunoprecipitated from both fractions, and bound products Western blotted for the presence of Ajuba. Snail was only detected in the nuclear extract and, therein, Snail associated with Ajuba (Fig. 2C). This, in combination with the immunofluorescence data, indicated that Ajuba accumulates in the nucleus upon expression of Snail and that it interacts with Snail in the nucleus.

Ajuba acts as a corepressor to Snail

A central function of Snail during EMT is transcriptional repression of epithelial cell-cell adhesive receptors, such as E-cadherin (Batlle et al. 2000; Cano et al. 2000). Since Ajuba LIM proteins interact with Snail in the nucleus, we sought to determine the functional relevance of this interaction by testing whether co-expression of Ajuba affects the ability of Snail to repress transcription downstream of the E-cadherin promoter. Transient luciferase reporter assays were performed using a construct in which luciferase expression is driven by a region of the human E-cadherin promoter that contains all three Snail-binding E-boxes. MCF-7 human breast cancer cells were transfected with the reporter construct along with Snail and Ajuba alone or in combination. Both Snail and Ajuba alone were able to repress transcription downstream of the E-cadherin promoter (Snail had a greater effect than Ajuba), but when co-expressed, there was increased repression over either alone, suggesting that Ajuba cooperates with Snail to repress E-cadherin transcription (Fig. 3A). The ability of Ajuba alone to repress transcription could be due to the trace levels of endogenous Snail in MCF-7 cells (data not shown). Alternatively, in addition to an interaction with Snail, Ajuba may contribute to repression through a Snail-independent manner as has been observed for the interaction between other Ajuba LIM proteins and other nuclear transcriptional regulators (Sharp et al. 2004).

Accumulated data on Ajuba function has revealed that the LIM and preLIM regions interact with distinct targets within common signaling pathways or cellular functions (Marie et al. 2003; Pratt et al. 2005). Since Ajuba acts as a Snail corepressor and the LIM region directs its interaction with Snail, this raises the possibility that the preLIM region is available to coordinate assembly of a chromatin repressor complex on promoters containing bound Snail. If this is true, then overexpression of the LIM region alone may act as a dominant inhibitor to block the

assembly of a repressor complex at a Snail regulated promoter by interfering with the interaction between endogenous full-length LIM protein and Snail. Indeed, when the Ajuba LIM region was overexpressed, it resulted in de-repression of transcription downstream of the E-cadherin promoter, even in presence of exogenous Snail (Fig. 3A). This indicated that the LIM region could act in a dominant negative manner to inhibit the co-repressor function of full-length Ajuba LIM proteins and, thus, block Snail-dependent repression.

The biochemical basis of the LIM region's potential "dominant inhibitory" function could be due to the LIM region displacing the endogenous full length Ajuba LIM protein bound to Snail or preventing full-length Ajuba LIM proteins from interacting with Snail. To test whether the Ajuba LIM region interferes with the Snail-Ajuba interaction, HEK293 cells were cotransfected with fixed and equal amounts of Flag-Ajuba and HA-Snail, and increasing amounts of myc-Ajuba LIM region. Snail was immunoprecipitated, and bound products Western blotted for the presence of Ajuba to determine if the interaction between full-length Ajuba and Snail decreased in the presence of excess Ajuba LIM region. Indeed as the amount of Ajuba LIM region transfected increased, the interaction of Snail with full-length Ajuba decreased (Fig. 3B). Together, these results suggested that interaction with full-length LIM proteins contributes to the ability of Snail to efficiently repress transcription.

For Ajuba to function as a Snail co-repressor, it should be present at the promoters of Snail repressed genes, such as E-cadherin, in a Snail-dependent manner. To determine if this was true, Ajuba chromatin immunoprecipitations (ChIP) were performed in the presence or absence of Snail, using the endogenous E-cadherin promoter as the target sequence. As current antiserum to Snail does not immunoprecipitate endogenous levels of Snail present in chromatin complexes, we were required to express Flag-tagged Snail so as to immunoprecipitate Snail with

anti-Flag antibodies. HEK293 cells stably expressing Flag-Snail or control empty vector were generated. Expression of Snail in these cells resulted in decreased E-cadherin levels, indicating that Snail is active (Fig. 3C). Nuclear chromatin preparations from these cells were immunoprecipitated with preimmune serum (control) or antibodies to Flag, Snail, or Ajuba and PCR performed for the E-cadherin promoter using primers that flanked the three E-boxes Snail binds. Snail was present on the E-cadherin promoter in cells overexpressing Flag-Snail as shown by either Flag or Snail ChIP (Fig. 3D, top). Importantly, endogenous Ajuba was detected on the E-cadherin promoter, but only in the presence of Snail (Fig. 3D, bottom). This result demonstrated that Ajuba was present on the promoter of a gene physiologically repressed by Snail and that its presence on the promoter was dependent upon the presence of Snail. In sum, these data indicated that Ajuba LIM proteins are functional Snail co-repressors.

Overexpression of Ajuba and LIMD1 affects neural crest development in Xenopus

The roles of Snail and Slug have been well-characterized in Xenopus neural crest development, a well-defined developmental EMT process. Overexpression of Snail or Slug early in Xenopus development leads to increased neural crest induction and enhanced migration of neural crest precursors during development (LaBonne and Bronner-Fraser 1998; Carl et al. 1999; LaBonne and Bronner-Fraser 2000; Aybar et al. 2003). Importantly, neural crest development is dependent on the repressor function of Snail and Slug (LaBonne and Bronner-Fraser 2000; Aybar et al. 2003). This, coupled with the ease of manipulation of Xenopus embryos, led us to utilize Xenopus neural crest development to determine whether the interaction between Ajuba LIM proteins and Snail proteins is biologically relevant *in vivo*.

We first tested the effect of overexpression of mouse Ajuba and LIMD1 on Xenopus neural crest development. We confirmed that mAjuba and mLIMD1 co-immunoprecipitated with XSnail (data not shown). One cell of the two-cell stage X. laevis embryo was injected with mAjuba or mLIMD1 mRNA. This introduces the mRNA into one bilateral half of the embryo while the other half serves as an internal control. Injection of XSnail or XSlug mRNA served as a positive control for induction of neural crest. Embryos were co-injected with β-galactosidase mRNA to track the injected regions through development. The embryos were fixed and X-Gal stained at appropriate stages and *in situ* hybridization was performed for markers of the neural plate (Sox2), non-neuronal ectoderm (Epidermal keratin, Epikeratin) and neural crest (Slug, Twist, and FoxD3). We detected a marked increase in the area of neural crest on the side of the embryos injected with each mRNA (Fig. 4A). XSlug *in situs* revealed increased staining on the injected side in 37.5% (n=24) and 65.9% (n=85) of embryos expressing mAjuba or mLIMD1, respectively, as compared with 78.6% (n=56) and 59.5% (n=42) of embryos that were injected with XSlug or XSnail, respectively (Fig. 4A). Similar results were obtained with other neural crest markers Twist and FoxD3 (Fig. 4A). We did not observe a significant change in Sox2 or Epikeratin staining in embryos injected with mAjuba or mLIMD1 mRNA (Fig. 4B) at stages 14 and 12, respectively, indicating that initial determination of neural and epidermal fates occurred normally. Therefore, mAjuba and mLIMD1 phenocopy the effects of XSnail and XSlug to increase neural crest area.

The LIM region blocks neural crest development in vivo

In another approach to test the role of Ajuba LIM proteins in neural crest development *in vivo*, we exploited the dominant inhibitory role of the LIM region (see Fig. 3) to block all

endogenous Xenopus Ajuba LIM protein from interacting with XSnail or XSlug. mRNA of the LIM region of either mAjuba or mLIMD1 was injected into one cell of the two-cell stage Xenopus embryo. Strikingly, expression of either LIM region reduced or eliminated Slug or Twist staining on the injected half of the embryo (Fig. 5A). Injection of mAjuba LIM region mRNA reduced Slug staining on the injected half in 52% of embryos (n=25), while injection of mLIMD1 LIM region mRNA reduced Slug staining in 56% of embryos (n=75) and Twist staining in 71.4% of embryos (n=21) (Fig. 5A). This was not due to a nonspecific effect (e.g. affecting cell survival or all cell fates), as the side of the embryos injected with LIM region mRNA still expressed other ectodermal markers Sox2 and Epikeratin, and developed into tadpoles without apparent defects at a gross morphological level (Supplemental Fig. 1).

To demonstrate specificity of the LIM region as a dominant inhibitor of neural crest development as regulated by endogenous Ajuba LIM proteins, increasing amounts of full-length mLIMD1 mRNA were co-injected into one cell of the two-cell Xenopus embryo in combination with a fixed amount of mLIMD1 LIM region mRNA. Without co-expression of full-length LIMD1, mLIMD1 LIM region caused a decrease in neural crest on the injected side in 56% of embryos (n=75). Co-injection of equal amounts of full-length mLIMD1 reduced this to 32% of embryos (n=46), while co-injection of twice the amount of full-length mLIMD1 gave only 5.5% embryos (n=54) with decreased neural crest. Instead, 65% of embryos (n=54) showed an increase in neural crest development on the injected side, consistent with mLIMD1 overexpression (Fig. 5B).

Ajuba LIM proteins specifically affect Snail-induced neural crest development in vivo

We have shown that overexpression of mAjuba LIM proteins enhanced Xenopus neural crest development, similar to overexpression of XSnail or XSlug. Moreover, preventing Xenopus Ajuba LIM proteins from binding to their LIM region targets (by expression of the LIM region alone) inhibited neural crest development. The ability of the Ajuba LIM region to inhibit neural crest development could be due to blocking endogenous LIM protein-Snail interactions or to blocking the interaction of the endogenous LIM protein(s) with other proteins that influence neural crest development. To test directly whether the LIM region blocked Snail regulated neural crest development specifically, we co-injected mLIMD1 LIM region mRNA with XSnail or XSlug mRNA. We found that overexpression of the LIM region was sufficient to block the expansion of neural crest normally seen upon overexpression of XSnail and XSlug. In fact, co-injection of the mLIMD1 LIM region with XSlug resulted in an inhibition of neural crest development in 90.5% (n=24) and 60% (n=15) of embryos as detected by Slug and Twist staining, respectively (Fig. 6). Similarly, co-injection of the LIM region with XSnail resulted in an inhibition of neural crest development in 91.7% (n=21) and 46.7% (n=15) of embryos as detected by Slug and Twist staining, respectively (Fig. 6). While this experiment does not entirely eliminate the possibility of affecting other targets of the LIM domains, these results did indicate that the interaction of XSnail or XSlug with full-length Xenopus LIM proteins was necessary for XSnail- or XSlug-induced neural crest development.

Two Xenopus laevis orthologs in the Ajuba subfamily

Expression of mAjuba and mLIMD1 mRNA in Xenopus embryos resulted in increased neural crest development, consistent with their potential roles as Snail family corepressors. We

sought to identify endogenous Xenopus LIM protein(s) in order to determine whether this is a physiological role for Ajuba LIM proteins during Xenopus development. An EST database search identified four putative orthologs, suggesting that the Xenopus families are less complex than the mammalian families. Two proteins were in the Ajuba subfamily, and had closest homology to mammalian LIMD1 and WTIP (Fig. 7). We cloned and characterized Xenopus LIMD1 (XLIMD1, GenBank Accession DQ913740). Like its mammalian orthologs, XLIMD1 has three C-terminal LIM domains and a preLIM region. In the LIM region, XLIMD1 has 83% identity to mLIMD1 and 66% identity to mAjuba at the amino acid level. The preLIM region of XLIMD1 shares only 29% identity to mLIMD1, but has 80% identity in the N-terminal 61 amino acids (Supplemental Fig. 2).

We next determined endogenous XLIMD1 expression during Xenopus development. Two distinct antisense probes to XLIMD1 mRNA were used for *in situ* hybridization and gave the same result. Sense probes served as controls for specificity (Fig. 8A, left). Low level, non-specific, or ubiquitous staining of XLIMD1 was seen in blastula and gastrula stage embryos, however, by stage 13.5, XLIMD1 was enriched throughout the neural plate (data not shown). In stage 17-19 embryos, XLIMD1 was enriched in the territory encompassing the anterior neural plate and the premigratory neural crest (Fig. 8A). By tailbud stages (Stage 21-24), there was strong expression of XLIMD1 in the migrating neural crest, as well as in the eye, brain, and otic placode (Fig. 8A). By comparison, Snail and Slug are expressed in the prospective neural crest in neurula stages and in migrating neural crest during and after stage 21 (Linker et al. 2000). Therefore XLIMD1 expression during Xenopus development largely overlaps with XSnail and XSlug. To confirm co-localization in developing neural crest, double *in situ* hybridization was performed for XLIMD1 and XSlug. Expression patterns indeed overlapped in regions of neural

crest in neurula and tailbud stage embryos (Fig. 8B). Therefore, XLIMD1 colocalized with XSnail and XSlug in areas of developing neural crest, consistent with a physiological role in the development of this tissue.

Finally, we tested whether XLIMD1 had the same effect as mammalian orthologs on neural crest development. First, we confirmed that XLIMD1 co-immunoprecipitated with XSnail (data not shown). We next injected XLIMD1 mRNA into one cell of the two-cell embryo. Injected embryos were fixed and *in situ* hybridization performed for neural crest markers Slug, Twist and FoxD3. Similar to mammalian LIM proteins, ectopic expression of XLIMD1 resulted in 42.5% (n=40), 22.4% (n=49), and 66.7% (n=21) of embryos with increased neural crest on the injected side as detected by Slug, Twist, and FoxD3 staining, respectively (Fig. 9). Therefore, overexpression of both murine and Xenopus Ajuba LIM proteins, like overexpression of XSnail or XSlug, resulted in an expansion of the neural crest territory.

Discussion

Here, we show that Ajuba and LIMD1 are functional corepressors of the vertebrate Snail family. Based on these results, we propose a model in which Ajuba LIM proteins bind via their LIM regions to Snail at promoters of Snail-regulated genes (Fig. 10). This interaction is mediated predominantly by the SNAG domain of Snail and the LIM region of the LIM proteins, leaving the PreLIM region of Ajuba LIM proteins available to facilitate assembly of the transcriptional repressor complex at these sites.

In this model, the LIM region alone is predicted to act in a dominant inhibitory manner by blocking endogenous full-length LIM proteins from binding to Snail, thereby reducing Snail-dependent repression. In support of this model, expression of excess LIM region alone prevents the interaction of full-length LIM protein with Snail (see Fig. 3B) and gives a dominant inhibitory phenotype, both in transient reporter assays (Fig. 3A) and *in vivo* during Xenopus neural crest development (Fig. 5A). Furthermore, the effect of the LIM region on Xenopus neural crest development can be rescued by addition of excess full length LIMD1 (Fig. 5B), suggesting specificity of the action of the LIM region. Together, these data suggest that the full-length LIM protein is necessary for Snail-dependent functions. It remains a possibility that the LIM region has additional targets in the nucleus, and we cannot rule out that some effects are through alternative targets. Nonetheless, we provide strong evidence that Snail is an important nuclear target of Ajuba and LIMD1, and that Ajuba LIM proteins affect Snail and Slug function in the nucleus. Indeed, the LIM region of LIMD1 blocked neural crest development even in the presence of exogenous Snail or Slug expression (Fig. 6).

All members of the Ajuba subfamily of LIM proteins interact with Snail family members (Fig. 1B) and serve redundant functions, as observed upon expression of either Ajuba or LIMD1

in Xenopus (Fig. 4A). Therefore, to determine a loss of function phenotype for these proteins, it is likely that the function of all members of the Ajuba LIM family will have to be blocked to observe a phenotype. The potential for compensation between Ajuba, LIMD1, and WTIP is demonstrated by the fact that we do not observe gross developmental pathologies in the single knockout mice of Ajuba or LIMD1 (Marie et al. 2003; Feng and Longmore, unpublished data). Here, in both mammalian cell lines and Xenopus embryos, we have used the LIM region as a dominant inhibitory protein to potentially block binding of all Ajuba LIM family members to the Snail proteins. The use of the LIM region, therefore, would prevent compensation that would be likely if only one LIM subfamily protein were lost. Similarly, others found that the two LIM proteins Zyxin and LPP served redundant functions in the recruitment of VASP to cadherin-based junctions. A dominant inhibitory mutant of either protein that failed to bind VASP could block the normal function of both Zyxin and LPP in stabilizing nascent epithelial junctions (Hansen and Beckerle 2006). It will be important to further explore genetically the extent of compensation between the Ajuba LIM family proteins both in mice and in lower organisms.

Because the LIM region alone acts in a dominant inhibitory manner, we hypothesize that the preLIM region is important for the repressor activity of Ajuba, perhaps through interactions with other repressor complex proteins. Previously, we have shown that the preLIM and LIM regions of Ajuba function as modular domains bringing together separate proteins to regulate a common pathway or cellular function. For example, in epithelial cells the Ajuba LIM region directs Ajuba to adherens junctions via binding to α-catenin, while the preLIM region interacts directly with F-Actin, providing a link between the junctions and the cytoskeleton (Marie et al. 2003). In fibroblasts, the LIM region directs Ajuba to focal adhesions while the preLIM region interacts with p130cas, allowing for localized activation of Rac during cell migration (Pratt et al.

2005). It will be critical to define the role of the preLIM region in Snail-mediated repression in future experiments.

While Ajuba and its family members do not directly bind DNA, it is becoming increasingly clear that through protein-protein interactions with known transcription factors many LIM domain containing proteins, including members of the Ajuba and Zyxin families, can regulate expression of genes in a variety of cells. For example, LIMD1 was recently identified as a corepressor of the retinoblastoma protein (Sharp et al. 2004), WTIP inhibits transcriptional activation by WT-1 (Srichai et al. 2004), LPP is a coactivator of PEA3, an ETS domain transcription factor (Guo et al. 2006), and TRIP6 enhances transactivation by v-Rel (Zhao et al. 1999). As we have observed for the interaction between Ajuba LIM proteins and Snail proteins, in most cases, the LIM region directs the interaction with the DNA binding protein. This then leaves the preLIM region available to potentially coordinate the assembly of repression or activation complexes at these sites.

Snail family members contribute to many biological and cellular processes including gastrulation and neural crest development as well as cell proliferation, survival, adhesion, and migration (Barrallo-Gimeno and Nieto 2005). Precise mechanisms of regulation of these different roles in various cell types and organisms remain unclear. Using Xenopus neural crest development as a model for Snail or Slug function, we show that Ajuba LIM proteins cooperate with Snail and Slug *in vivo*. Overexpression of Ajuba LIM proteins gives increased neural crest, similar to overexpression of XSnail or XSlug (Fig. 4A). This role in Xenopus is consistent with Ajuba and LIMD1 acting as corepressors of Snail and Slug as the repressor activity of Snail and Slug has been shown to be essential for their roles in neural crest development (LaBonne and Bronner-Fraser 2000; Aybar et al. 2003). Importantly, blocking the interaction of full-length

Ajuba LIM proteins with Snail by expressing the Ajuba or LIMD1 LIM regions alone blocks neural crest development (Fig. 6). This also results in a loss of at least one neural crest derivative, melanocytes, at later stages (see Supplemental Fig. 1). The genes directly repressed by Snail and Slug in Xenopus are unknown, but as these are uncovered, it will be important to determine whether Ajuba LIM proteins affect all, or a subset, of these genes. In addition, it will be interesting to expand these studies to determine the extent to which Ajuba LIM proteins affect Snail function in other processes.

Here, we describe a nuclear role for Ajuba LIM proteins as corepressors with Snail that contribute to repression of E-cadherin. Previously, Ajuba, LIMD1, and WTIP all have been shown to localize to cell-cell adhesive junctions in epithelial cells (Marie et al. 2003; Srichai et al. 2004). Ajuba contributes to adherens junction formation and/or stability, in part by coupling the E-cadherin adhesive complex to the actin cytoskeleton (Marie et al. 2003). Because of these distinct roles for the separate pools of Ajuba, it may play an integral role in communicating from adhesive complexes to the nucleus to provide precise regulation of this process. While we do not have evidence that Ajuba LIM proteins can initiate EMT, we do show that expression of Snail can result in the trapping of Ajuba in the nucleus (Fig. 2A and 2B). Once in the nucleus, Ajuba LIM proteins can act as co-repressors to Snail family members, and could contribute to a feed forward loop to maintain the mesenchymal phenotype. By contributing to repression of E-cadherin, Ajuba may also indirectly allow more Ajuba to be available for entry into the nucleus (as Ajuba is normally present in E-cadherin dependent adherens junctions). Consistent with this, Jamora, et al. found that Ajuba is released from adherens junctions upon expression of Snail and subsequent transcriptional downregulation of E-cadherin expression (Jamora et al. 2005). We have also observed the release of Ajuba from epithelial junctions in HaCaT epithelial cells upon

treatment with TGF-β, which induces EMT (Langer and Longmore, unpublished data). In sum, Ajuba LIM proteins could play an important role in the regulation of EMT by providing communication between adherens junctions and the nucleus during this dynamic transition.

It is also of interest to determine whether Ajuba LIM proteins play a role in the reverse process, a mesenchymal to epithelial transition (MET). MET occurs physiologically during mammalian nephrogenesis (Kanwar et al. 2004) and pathologically upon redifferentiation of metastatic cells as they reform epithelial-like tumors at metastatic sites (Thiery and Sleeman 2006). Recently, Snail was shown to be capable of repressing its own transcription (Peiro et al. 2006), perhaps providing a mechanism to initiate MET. As Snail expression is decreased, Ajuba could exit the nucleus, be recruited to newly forming adherens junctions, and thereby contribute to epithelia formation.

Materials and Methods

Cell Culture, Transfection, and Antibodies

HEK293T, MCF-7, HaCaT, and MDA-231 cells were grown in DMEM supplemented with 10% FBS, 2mM L-Glutamine, 100 units/ml penicillin and 100μg/ml streptomycin. Stable cell lines of HaCaT expressing myc-Ajuba were selected and grown in G418 and HEK293T cell expressing Snail were selected and grown in Zeocin. All transfections were performed with Trans-IT LT-1 (Mirus) according to the manufacturer's protocol. Antibodies used were: Ajuba (Cell Signaling, and (Pratt et al. 2005)), Snail (Santa Cruz), and HA (Sigma). HRP conjugated antibodies to Flag and Myc tags were from Sigma.

Plasmids

pCS2-myc-mAjuba, pCS2-myc-mLIMD1, pCS2-myc-LPP, pCS2-myc-Zyxin, pCS2-myc-mAjubaPreLIM, pCS2-myc-mAjubaLIM, and pcDNA3.1-RFP-mAjuba have been described previously (Goyal et al. 1999; Marie et al. 2003; Pratt et al. 2005). pCMV14-mSnail-Flag was a kind gift from H. L. Grimes (University of Louisville). pCMV5-HA-mSnail full-length and ΔSNAG constructs were generated through PCR into the pCMV5-HA vector. pcDNA3.1-GFP-hSnailWT and pcDNA3.1-GFP-hSnail8SA were subcloned into pcDNA3.1 (Invitrogen) from the corresponding CMV-Tag constructs, which were from M. Hung (University of Texas M.D. Anderson Cancer Center). pCS2-XSlug was from C. LaBonne (Northwestern University) and pCS2-myc-XSnail was constructed by PCR from an X. laevis cDNA library. All subcloning was confirmed by DNA sequencing.

Immunoprecipitation, Immunoblots, and Immunofluorescence

Cells were lysed in IP buffer (10mM Tris pH7.4, 150mM NaCl, 1mM EDTA 10mM NaF, 10% glycerol, 1% NP-40, protease inhibitors) and extract clarified by spinning at

12,000rpm for 20 minutes. Lysates were pre-cleared with protein A or G beads alone for 1 hour, then incubated overnight with antibody. Protein A or G beads were added for 1 hour, washed 4 times with IP buffer, and bound proteins eluted directly into SDS-PAGE sample buffer. For Flag M2 conjugated beads (Sigma) were used for immunoprecipitation and bound proteins eluted by competition with Flag peptide. SDS-PAGE and immunoblots were performed using standard protocols. For nuclear and cytosolic extracts, HaCaT cells were lysed in a hypotonic buffer, cell membranes disrupted by dounce homogenization, and nuclei pelleted by centrifugation. Supernatant was used as cytosolic extract. Nuclei were then extracted in hypertonic lysis buffer and centrifuged again. Extracts were adjusted to isotonic, and immunoprecipitations performed.

Immunofluorescence was performed as previously described (Marie et al. 2003). Images were taken on a Zeiss confocal microscope using the LSM 550 software. Localization of Ajuba was quantified by counting at least 100 cells co-transfected with GFP and RFP constructs.

Luciferase Assay

MCF-7 cells were transiently transfected with an 0.12μg E-cadherin-luciferase reporter construct (Hajra et al. 2002), 0.12μg CMV-β-galactosidase and combinations of empty vector, mAjuba, mAjuba LIM, and mSnail to equal 1.2μg total DNA. Cells were lysed in Cell Culture Lysis Reagent (Promega), and lysate used in both a luminescent β-galactosidase detection assay (BD Biosciences) and luciferase assay. All samples were read on a luminometer and values shown were obtained by normalizing luciferase values to β-galactosidase activity.

Chromatin Immunoprecipitation

Cells were grown in 150mm plates to 70-90% confluency, fixed in 1% formaldehyde, and harvested, and chromatin immunoprecipitations performed using the EZ-CHIP kit (Upstate) according to the manufacturer's protocol. PCR was performed on immunoprecipitated DNA

using the following primer pairs: 5' AATCAGAACCGTGCAGGTCC 3' and 5' ACAGGTGCTTTGCAGTTCCG 3' (see Figure 3D)

X. laevis embryos and mRNA injections

X. laevis albino embryos were obtained through *in vitro* fertilization and raised as described previously (Kroll et al. 1998). Embryos were staged following Nieuwkoop and Faber (Nieuwkoop and Faber, 1967). For injections, capped mRNA was transcribed *in vitro* using the mMessage mMachine kit (Ambion). RNAs (250-500pg) were coinjected into one blastomere of two-cell stage embryos with β-galactosidase mRNA (50pg).

In situ hybridizations

Embryos were raised until indicated stages, fixed for 1 hour in MEMFA, X-gal (5-bromo-4-chloro-3-indolyl β-galactopyranoside) stained and analyzed by *in situ* hybridizations as described (Harland 1991). For XLIMD1 *in situs*, two IMAGE clones of XLIMD1, 5513507 (CF271778) and 6640526 (BU914068), that overlap the preLIM region (nts 1-573) and LIM region (nts 1330-1836) were used. Sense and antisense *in situ* probes of each XLIMD1 clone and antisense probes of Sox2 (Mizuseki et al. 1998), Epidermal keratin (Epikeratin) (Jonas et al. 1985), Slug (Mayor et al. 1995), Twist (Hopwood et al. 1989), and FoxD3 (Bellmeyer et al. 2003) were generated by *in vitro* transcription with digoxigenin-11-UTP (Roche). Probes were detected with alkaline phosphatase (AP)-conjugated anti-digoxigenin antibodies (Roche) with Nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl-phosphate (NBT/BCIP; Roche). For double *in situs*, the Dig-labeled XLIMD1 probe (6640526) was used in combination with a probe to Slug generated by *in vitro* transcription with fluorescein-12-UTP (Roche). Probes were detected using AP-anti-digoxenin (Roche) with BCIP (Roche) and AP-anti-fluorescein (Roche) with Magenta phosphate (5-bromo-6-chloro-3-indoxyl phosphate, p-toluidine salt; Biosynth).

Identification of X. laevis orthologs and cloning of X. laevis LIMD1

An EST database search revealed partial sequences overlapping the LIM regions of Zyxin (BU907427, BG234393, BJ068010, BI349606, and BG409548), LPP (BC077865, BU903395, CB983325, BG363738, BU912724, BM179756, and BG364767), LIMD1 (BU914068 and AW642119), and WTIP (BJ068924 and AW645564). Phylogenetic relationships between the LIM regions of M. musculus, D. melanogaster, C. elegans, and putative X. laevis LIM proteins were generated using the Clustal W algorithm and MegAlign software (DNASTAR). The full LIM region was used for all proteins except X. laevis WTIP, where only two of three LIM domains have been identified. Gene-specific primers to full length XLIMD1 were designed based on available EST sequences (CF271778 and BU914068). With these, the full-length XLIMD1 cDNA was amplified from a X. laevis stage 24 embryo cDNA library by PCR. This fragment was cloned into pCS2 and sequenced (Gen Bank Accession DQ913740).

Acknowledgments

EML is an HHMI predoctoral fellow. This work was supported by grants NIH GM66815-01, the March of Dimes (#1-FY06-374), and the American Cancer Society (RSG-06-148-01-DDC) to KLK; NIH Core CA10815, CA92088, CA095561, DAMD17-96-1-6141, 17-02-1-0631, the Irving A. Hansen Memorial Foundation, the Susan G. Komen Breast Cancer Foundation, and the Emerald Foundation to FJR; and NIH CA75315, CA106496, and the Washington University/Pfizer Biomedical Research Program to GDL. GDL was an Established Investigator of the American Heart Association.

References

Aybar, M.J., Nieto, M.A., and Mayor, R. 2003. Snail precedes slug in the genetic cascade required for the specification and migration of the Xenopus neural crest. *Development* **130**(3): 483-494.

Barrallo-Gimeno, A. and Nieto, M.A. 2005. The Snail genes as inducers of cell movement and survival: implications in development and cancer. *Development* **132**(14): 3151-3161.

Batlle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J., and Garcia De Herreros, A. 2000. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol* **2**(2): 84-89.

Bellmeyer, A., Krase, J., Lindgren, J., and LaBonne, C. 2003. The protooncogene c-myc is an essential regulator of neural crest formation in xenopus. *Dev Cell* **4**(6): 827-839.

Blanco, M.J., Moreno-Bueno, G., Sarrio, D., Locascio, A., Cano, A., Palacios, J., and Nieto, M.A. 2002. Correlation of Snail expression with histological grade and lymph node status in breast carcinomas. *Oncogene* **21**(20): 3241-3246.

Cano, A., Perez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. 2000. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol* **2**(2): 76-83.

Carl, T.F., Dufton, C., Hanken, J., and Klymkowsky, M.W. 1999. Inhibition of neural crest migration in Xenopus using antisense slug RNA. *Dev Biol* **213**(1): 101-115.

Carver, E.A., Jiang, R., Lan, Y., Oram, K.F., and Gridley, T. 2001. The mouse snail gene encodes a key regulator of the epithelial-mesenchymal transition. *Mol Cell Biol* **21**(23): 8184-8188.

Comijn, J., Berx, G., Vermassen, P., Verschueren, K., van Grunsven, L., Bruyneel, E., Mareel, M., Huylebroeck, D., and van Roy, F. 2001. The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol Cell* **7**(6): 1267-1278.

del Barrio, M.G. and Nieto, M.A. 2002. Overexpression of Snail family members highlights their ability to promote chick neural crest formation. *Development* **129**(7): 1583-1593.

Dominguez, D., Montserrat-Sentis, B., Virgos-Soler, A., Guaita, S., Grueso, J., Porta, M., Puig, I., Baulida, J., Franci, C., and Garcia de Herreros, A. 2003. Phosphorylation regulates the subcellular location and activity of the snail transcriptional repressor. *Mol Cell Biol* **23**(14): 5078-5089.

Goyal, R.K., Lin, P., Kanungo, J., Payne, A.S., Muslin, A.J., and Longmore, G.D. 1999. Ajuba, a novel LIM protein, interacts with Grb2, augments mitogen-activated protein kinase activity in fibroblasts, and promotes meiotic maturation of Xenopus oocytes in a Grb2- and Ras-dependent manner. *Mol Cell Biol* **19**(6): 4379-4389.

Grimes, H.L., Chan, T.O., Zweidler-McKay, P.A., Tong, B., and Tsichlis, P.N. 1996. The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. *Mol Cell Biol* **16**(11): 6263-6272.

Guaita, S., Puig, I., Franci, C., Garrido, M., Dominguez, D., Batlle, E., Sancho, E., Dedhar, S., De Herreros, A.G., and Baulida, J. 2002. Snail induction of epithelial to mesenchymal transition in tumor cells is accompanied by MUC1 repression and ZEB1 expression. *J Biol Chem* **277**(42): 39209-39216.

Guo, B., Sallis, R.E., Greenall, A., Petit, M.M., Jansen, E., Young, L., Van de Ven, W.J., and Sharrocks, A.D. 2006. The LIM domain protein LPP is a coactivator for the ETS domain transcription factor PEA3. *Mol Cell Biol* **26**(12): 4529-4538.

Hajra, K.M., Chen, D.Y., and Fearon, E.R. 2002. The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res* **62**(6): 1613-1618.

Hansen, M.D. and Beckerle, M.C. 2006. Opposing roles of zyxin/LPP ACTA repeats and the LIM domain region in cell-cell adhesion. *J Biol Chem* **281**(23): 16178-16188.

Harland, R.M. 1991. In situ hybridization: an improved whole-mount method for Xenopus embryos. *Methods Cell Biol* **36**: 685-695.

Hemavathy, K., Guru, S.C., Harris, J., Chen, J.D., and Ip, Y.T. 2000. Human Slug is a repressor that localizes to sites of active transcription. *Mol Cell Biol* **20**(14): 5087-5095.

Hemavathy, K., Hu, X., Ashraf, S.I., Small, S.J., and Ip, Y.T. 2004. The repressor function of snail is required for Drosophila gastrulation and is not replaceable by Escargot or Worniu. *Dev Biol* **269**(2): 411-420.

Hoffman, L.M., Jensen, C.C., Kloeker, S., Wang, C.L., Yoshigi, M., and Beckerle, M.C. 2006. Genetic ablation of zyxin causes Mena/VASP mislocalization, increased motility, and deficits in actin remodeling. *J Cell Biol* **172**(5): 771-782.

Hopwood, N.D., Pluck, A., and Gurdon, J.B. 1989. A Xenopus mRNA related to Drosophila twist is expressed in response to induction in the mesoderm and the neural crest. *Cell* **59**(5): 893-903.

Ikenouchi, J., Matsuda, M., Furuse, M., and Tsukita, S. 2003. Regulation of tight junctions during the epithelium-mesenchyme transition: direct repression of the gene expression of claudins/occludin by Snail. *J Cell Sci* **116**(Pt 10): 1959-1967.

Jamora, C., Lee, P., Kocieniewski, P., Azhar, M., Hosokawa, R., Chai, Y., and Fuchs, E. 2005. A signaling pathway involving TGF-beta2 and snail in hair follicle morphogenesis. *PLoS Biol* **3**(1): e11.

Jonas, E., Sargent, T.D., and Dawid, I.B. 1985. Epidermal keratin gene expressed in embryos of Xenopus laevis. *Proc Natl Acad Sci U S A* **82**(16): 5413-5417.

Kajita, M., McClinic, K.N., and Wade, P.A. 2004. Aberrant expression of the transcription factors snail and slug alters the response to genotoxic stress. *Mol Cell Biol* **24**(17): 7559-7566.

Kalluri, R. and Neilson, E.G. 2003. Epithelial-mesenchymal transition and its implications for fibrosis. *J Clin Invest* **112**(12): 1776-1784.

Kanungo, J., Pratt, S.J., Marie, H., and Longmore, G.D. 2000. Ajuba, a cytosolic LIM protein, shuttles into the nucleus and affects embryonal cell proliferation and fate decisions. *Mol Biol Cell* **11**(10): 3299-3313.

Kanwar, Y.S., Wada, J., Lin, S., Danesh, F.R., Chugh, S.S., Yang, Q., Banerjee, T., and Lomasney, J.W. 2004. Update of extracellular matrix, its receptors, and cell adhesion molecules in mammalian nephrogenesis. *Am J Physiol Renal Physiol* **286**(2): F202-215.

Kroll, K.L., Salic, A.N., Evans, L.M., and Kirschner, M.W. 1998. Geminin, a neuralizing molecule that demarcates the future neural plate at the onset of gastrulation. *Development* **125**(16): 3247-3258.

LaBonne, C. and Bronner-Fraser, M. 1998. Neural crest induction in Xenopus: evidence for a two-signal model. *Development* **125**(13): 2403-2414.

-. 2000. Snail-related transcriptional repressors are required in Xenopus for both the induction of the neural crest and its subsequent migration. *Dev Biol* **221**(1): 195-205.

Linker, C., Bronner-Fraser, M., and Mayor, R. 2000. Relationship between gene expression domains of Xsnail, Xslug, and Xtwist and cell movement in the prospective neural crest of Xenopus. *Dev Biol* **224**(2): 215-225.

Marie, H., Pratt, S.J., Betson, M., Epple, H., Kittler, J.T., Meek, L., Moss, S.J., Troyanovsky, S., Attwell, D., Longmore, G.D., and Braga, V.M. 2003. The LIM protein Ajuba is recruited to cadherin-dependent cell junctions through an association with alpha-catenin. *J Biol Chem* **278**(2): 1220-1228.

Mayor, R., Morgan, R., and Sargent, M.G. 1995. Induction of the prospective neural crest of Xenopus. *Development* **121**(3): 767-777.

Mizuseki, K., Kishi, M., Matsui, M., Nakanishi, S., and Sasai, Y. 1998. Xenopus Zic-related-1 and Sox-2, two factors induced by chordin, have distinct activities in the initiation of neural induction. *Development* **125**(4): 579-587.

Moody, S.E., Perez, D., Pan, T.C., Sarkisian, C.J., Portocarrero, C.P., Sterner, C.J., Notorfrancesco, K.L., Cardiff, R.D., and Chodosh, L.A. 2005. The transcriptional repressor Snail promotes mammary tumor recurrence. *Cancer Cell* **8**(3): 197-209.

Nakayama, H., Scott, I.C., and Cross, J.C. 1998. The transition to endoreduplication in trophoblast giant cells is regulated by the mSNA zinc finger transcription factor. *Dev Biol* **199**(1): 150-163.

Nibu, Y., Zhang, H., Bajor, E., Barolo, S., Small, S., and Levine, M. 1998. dCtBP mediates transcriptional repression by Knirps, Kruppel and Snail in the Drosophila embryo. *Embo J* **17**(23): 7009-7020.

Nieto, M.A., Sargent, M.G., Wilkinson, D.G., and Cooke, J. 1994. Control of cell behavior during vertebrate development by Slug, a zinc finger gene. *Science* **264**(5160): 835-839.

Nieuwkoop, P.D., Faber, J. 1967. *Normal Table of Xenopus laevis*. North Holland Publishing Company, Amsterdam.

Nix, D.A. and Beckerle, M.C. 1997. Nuclear-cytoplasmic shuttling of the focal contact protein, zyxin: a potential mechanism for communication between sites of cell adhesion and the nucleus. *J Cell Biol* **138**(5): 1139-1147.

Peinado, H., Ballestar, E., Esteller, M., and Cano, A. 2004. Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex. *Mol Cell Biol* **24**(1): 306-319.

Peiro, S., Escriva, M., Puig, I., Barbera, M.J., Dave, N., Herranz, N., Larriba, M.J., Takkunen, M., Franci, C., Munoz, A., Virtanen, I., Baulida, J., and Garcia de Herreros, A. 2006. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic Acids Res* **34**(7): 2077-2084.

Perez-Moreno, M.A., Locascio, A., Rodrigo, I., Dhondt, G., Portillo, F., Nieto, M.A., and Cano, A. 2001. A new role for E12/E47 in the repression of E-cadherin expression and epithelial-mesenchymal transitions. *J Biol Chem* **276**(29): 27424-27431.

Pratt, S.J., Epple, H., Ward, M., Feng, Y., Braga, V.M., and Longmore, G.D. 2005. The LIM protein Ajuba influences p130Cas localization and Rac1 activity during cell migration. *J Cell Biol* **168**(5): 813-824.

Savagner, P., Kusewitt, D.F., Carver, E.A., Magnino, F., Choi, C., Gridley, T., and Hudson, L.G. 2005. Developmental transcription factor slug is required for effective re-epithelialization by adult keratinocytes. *J Cell Physiol* **202**(3): 858-866.

Sharp, T.V., Munoz, F., Bourboulia, D., Presneau, N., Darai, E., Wang, H.W., Cannon, M., Butcher, D.N., Nicholson, A.G., Klein, G., Imreh, S., and Boshoff, C. 2004. LIM domains-containing protein 1 (LIMD1), a tumor suppressor encoded at chromosome

3p21.3, binds pRB and represses E2F-driven transcription. *Proc Natl Acad Sci U S A* **101**(47): 16531-16536.

Srichai, M.B., Konieczkowski, M., Padiyar, A., Konieczkowski, D.J., Mukherjee, A., Hayden, P.S., Kamat, S., El-Meanawy, M.A., Khan, S., Mundel, P., Lee, S.B., Bruggeman, L.A., Schelling, J.R., and Sedor, J.R. 2004. A WT1 co-regulator controls podocyte phenotype by shuttling between adhesion structures and nucleus. *J Biol Chem* **279**(14): 14398-14408.

Thiery, J.P. 2002. Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer* **2**(6): 442-454.

Thiery, J.P. and Sleeman, J.P. 2006. Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol* **7**(2): 131-142.

Vega, S., Morales, A.V., Ocana, O.H., Valdes, F., Fabregat, I., and Nieto, M.A. 2004. Snail blocks the cell cycle and confers resistance to cell death. *Genes Dev* **18**(10): 1131-1143.

Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R.A. 2004. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117**(7): 927-939.

Zhao, M.K., Wang, Y., Murphy, K., Yi, J., Beckerle, M.C., and Gilmore, T.D. 1999. LIM domain-containing protein trip6 can act as a coactivator for the v-Rel transcription factor. *Gene Expr* **8**(4): 207-217.

Zhou, B.P., Deng, J., Xia, W., Xu, J., Li, Y.M., Gunduz, M., and Hung, M.C. 2004. Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. *Nat Cell Biol* **6**(10): 931-940.

Figure Legends

Figure 1. Ajuba LIM proteins interact with the Snail family of transcriptional repressors. A. Myc-tagged LIM proteins and Flag-tagged Snail were cotransfected into HEK293 cells. Snail was immunoprecipitated from lysates (anti-Flag) and bound products Western blotted for LIM protein (anti-myc) and Snail (anti-Flag). Control Western blot of lysate is on right panel of each set. B. Table summarizing the interaction between Snail family members and Ajuba or Zyxin family members, as determined using the method described in (A). ND – not determined. C. Top: Schematic diagram of Ajuba constructs used. NES – nuclear export sequence. Bottom: Co-immunoprecipitation experiments as described for (A). D. Top: Schematic diagram of full length and mutant Snail constructs used. Bottom: Myc-tagged Ajuba and HA-tagged Snail constructs were cotransfected into HEK293 cells. Snail was immunoprecipitated (anti-HA) and bound products Western blotted for Ajuba (anti-myc) and Snail (anti-HA). Control Western blot of lysate is shown on the left. The amount of Ajuba immunoprecipitated relative to input was quantified by densitometry and controlled for the amount of Snail immunoprecipitated. (Ajuba immunoprecipitated with full-length Snail was arbitrarily set to equal 1). E. Endogenous Snail was immunoprecipitated from lysates of HaCaT or MDA-231 cells and bound products Western blotted for the presence of Ajuba, LIMD1 and Snail. Controls include pulldown with Protein G beads alone and lysate input.

Figure 2. Ajuba interacts with Snail in the nucleus. A. Confocal immunofluorescence analysis of MCF-7 breast cancer cells transfected with RFP-Ajuba in combination with GFP alone or GFP-Snail. Arrows indicate cotransfected cells. B. Quantification of immunofluorescence results. Shown are the percent of cells in which RFP-Ajuba is localized to

the cytosol (white bars) or nucleus (black bars) in the presence of GFP, GFP-SnailWT or GFP-Snail8SA. For each sample, at least 100 cells were counted. The experiment was repeated three times with similar results. Shown is one representative experiment. C. Myc-Ajuba was stably overexpressed in HaCaT epithelial cells. Endogenous Snail was immunoprecipitated from nuclear extracts (NE) or cytosolic extracts (CE) of these cells and bound products Western blotted for the presence of Ajuba (anti-myc) and Snail. Control immunoprecipitation was performed with rabbit preimmune sera (PI). Western blot of input controls is shown on the left.

Figure 3. Ajuba is a corepressor of Snail. A. Transient luciferase reporter assay using luciferase driven by the human E-cadherin promoter. Constructs, as indicated, were expressed in MCF-7 cells and luciferase activity determined and normalized to β-galactosidase activity (from co-transfected CMV-β−galactosidase construct). Experiments were performed in triplicate. Shown are mean normalized luciferase values +/- standard deviation. B. HEK293 cells were co-transfected with Flag-Ajuba, HA-Snail, and myc-Ajuba LIM region as indicated. Snail was immunoprecipitated from lysates (anti-HA) and bound products Western blotted for full-length Ajuba (anti-Flag), Snail (anti-HA), and Ajuba LIM region (anti-myc). The amount of Ajuba immunoprecipitated relative to input was quantified by densitometry and controlled for the amount of Snail immunoprecipitated. The value obtained for lane 4 was arbitrarily set to equal 1. C. HEK293 cells were stably transfected with vector or Flag-Snail. Lysates of cells were immunoblotted for E-cadherin, Flag, and Ajuba. Tubulin is a loading control. D. ChIPs were performed in HEK293 cells stably transfected for vector or Flag-Snail with antibodies to Flag or Snail (top panel) and Ajuba (bottom panel). Preimmune (PI) sera was used as a control. PCRs

were performed using primers flanking the three E-boxes in the E-cadherin promoter. In schematic, A, B, and C represent E-boxes on the E-cadherin promoter.

Figure 4. Ajuba and LIMD1 specifically affect neural crest development in Xenopus. A. X. laevis embryos were co-injected with β-gal and either mAjuba, mLIMD1, XSlug, or XSnail capped mRNAs in one cell of the two-cell stage embryo. Embryos were fixed at stage 18-19 and *in situ* hybridization performed for XSlug, XTwist, or XFoxD3. Shown are dorsal views with anterior end oriented up. X-Gal staining shows injected region of the embryo (oriented to the right). Numbers in the lower right-hand corner indicate the number of embryos that display the phenotype shown over the number of embryos injected. B. X. laevis embryos were co-injected with β-gal and either mAjuba or mLIMD1 and fixed at stage 12 (top) and 14 (bottom) and *in situ* hybridization performed for Epikeratin and Sox2, respectively. Top panels are in side view with blastopore oriented left and bottom panels are dorsal views with anterior oriented up. X-Gal staining shows injected region of the embryo.

Figure 5. The LIM region of Ajuba and LIMD1 block neural crest development in Xenopus. A. X. laevis embryos were co-injected with β-gal and either mAjuba LIM region or mLIMD1 LIM region capped mRNA, fixed at stage 18-19, and *in situ* hybridized for XSlug and XTwist. Shown are dorsal views with anterior oriented up. X-Gal staining shows injected region of the embryo (oriented to the right). Numbers in the lower right-hand corner show the number of embryos that display the phenotype shown over the number of embryos injected. ND – not determined. B. X. laevis embryos were co-injected with β-gal and combinations of mLIMD1 LIM region +/- full-length mLIMD1 as shown. Embryos were fixed at stage 18-19

and *in situ* hybridization performed for Slug. Graph shows percent of embryos with decreased Slug (black bars) or increased Slug (white bars) staining on the injected half of the embryo.

Figure 6. The LIM region of Ajuba or LIMD1 specifically blocks Snail- or Slug-induced neural crest development in Xenopus. X. laevis embryos were co-injected with β-gal, mLIMD1 LIM region, and either XSlug or XSnail capped mRNA. Embryos were fixed at stage 18-19 and *in situ* hybridization performed for XSlug and XTwist. Shown are dorsal views with anterior oriented up. X-Gal staining shows injected region of the embryo (oriented to the right). Numbers in the lower right-hand corner show the number of embryos that display the phenotype shown over the number of embryos injected.

Figure 7. X. laevis orthologs of the Ajuba and Zyxin families of LIM proteins. A predicted phylogenetic relationship of the LIM regions of the C. elegans, Drosophila, mouse and putative X. laevis orthologs of the Ajuba/Zyxin family of LIM proteins. *For X. laevis WTIP, only the first two LIM domains of the LIM region have been identified.

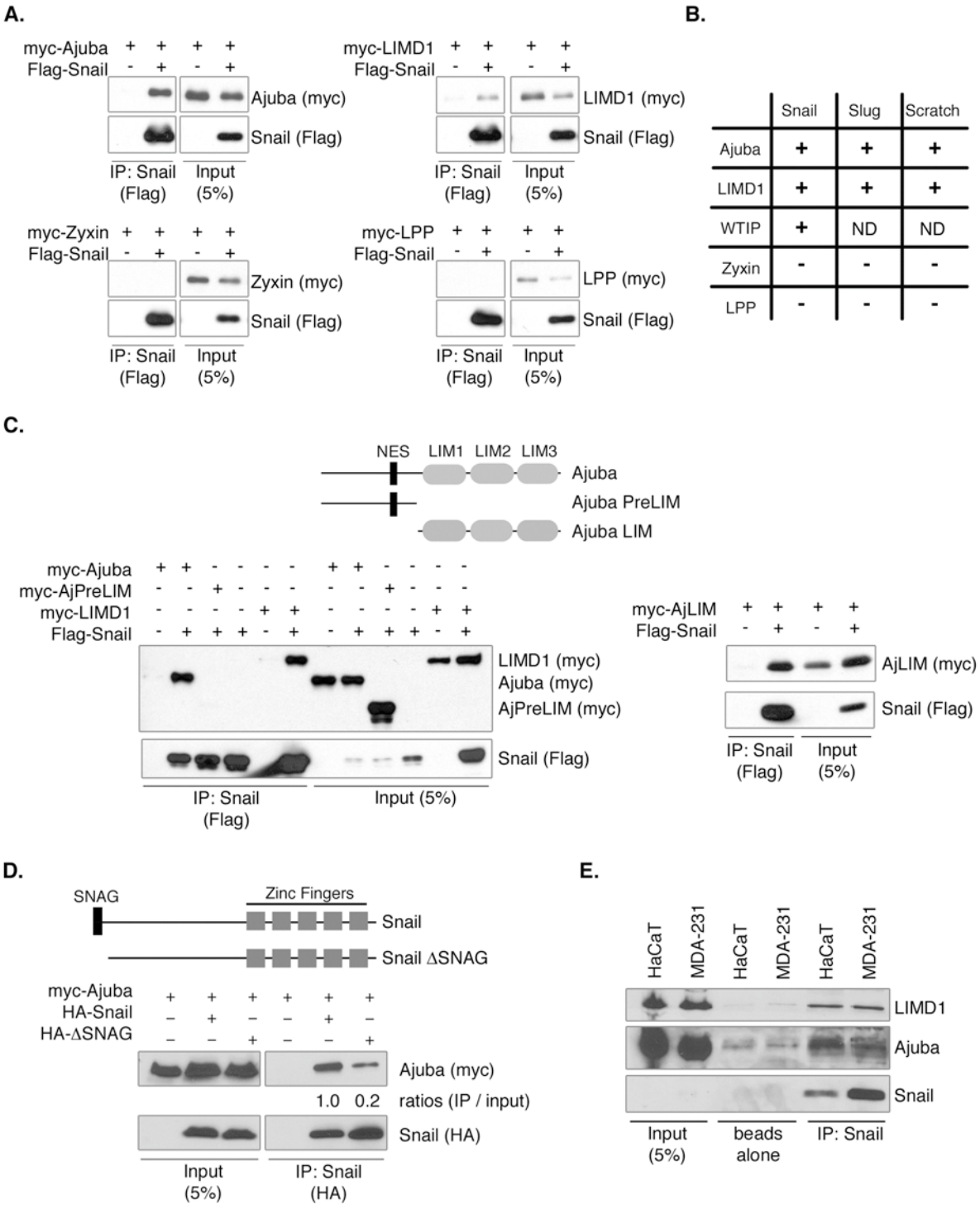
Figure 8. Endogenous XLIMD1 expression pattern overlaps that of XSlug. A. *In situ* hybridization was performed on uninjected X. laevis embryos using sense and antisense probes for XLIMD1. Staining patterns at stages 19, 21, and 24 are shown. Arrowheads indicate XLIMD1 staining in the premigratory (Stage 19) and migratory (Stages 21 and 24) neural crest. Embryos are oriented with dorsal up and anterior to the right. B. Double *in situ* hybridization for XLIMD1 and XSlug in X. laevis embryos at the stages shown. XLIMD1 localization alone is shown in blue (left column) and XSlug is overlayed in pink on right. Arrows indicate areas of

colocalization in neural crest derivatives. Embryos are oriented with dorsal up and anterior to the right.

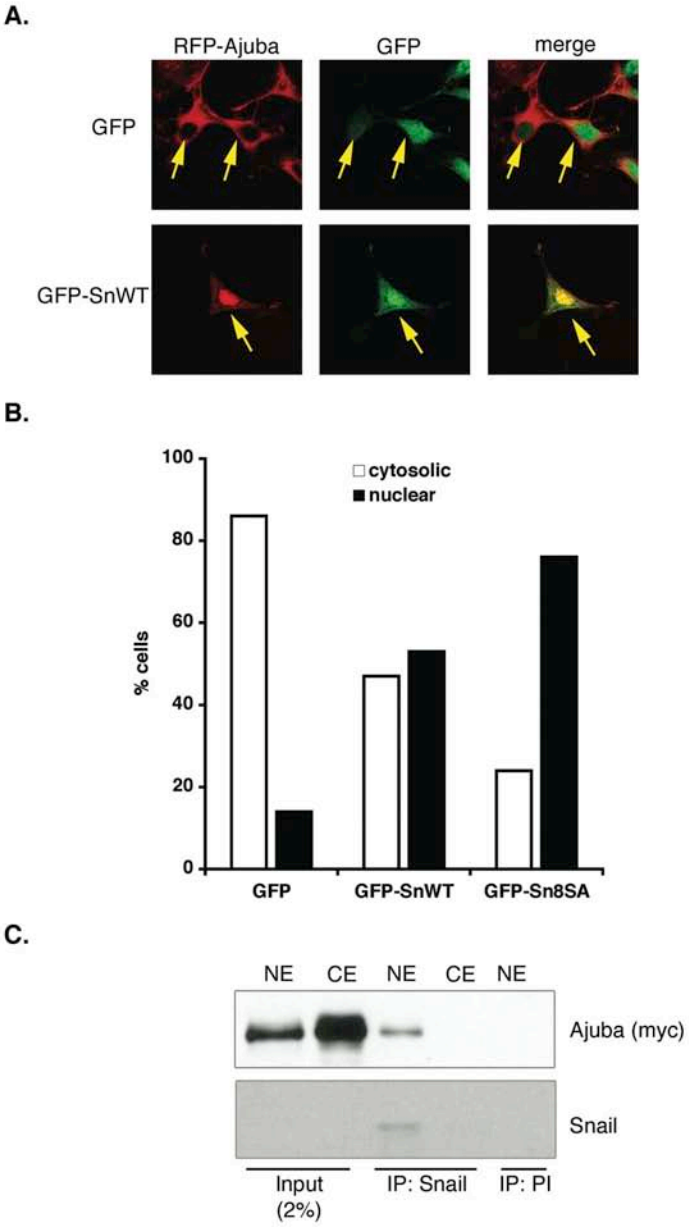
Figure 9. XLIMD1 expression results in an increase in neural crest development. X. laevis embryos were co-injected with β-gal and XLIMD1 capped mRNA in one cell of the two-cell stage embryo. Embryos were fixed at stage 18-19 and *in situ* hybridization performed for XSlug, XTwist, or XFoxD3. X-Gal staining (oriented to the right) shows the injected region of the embryo. Numbers in the lower right-hand corner indicate the number of embryos that display the phenotype shown over the number of embryos injected.

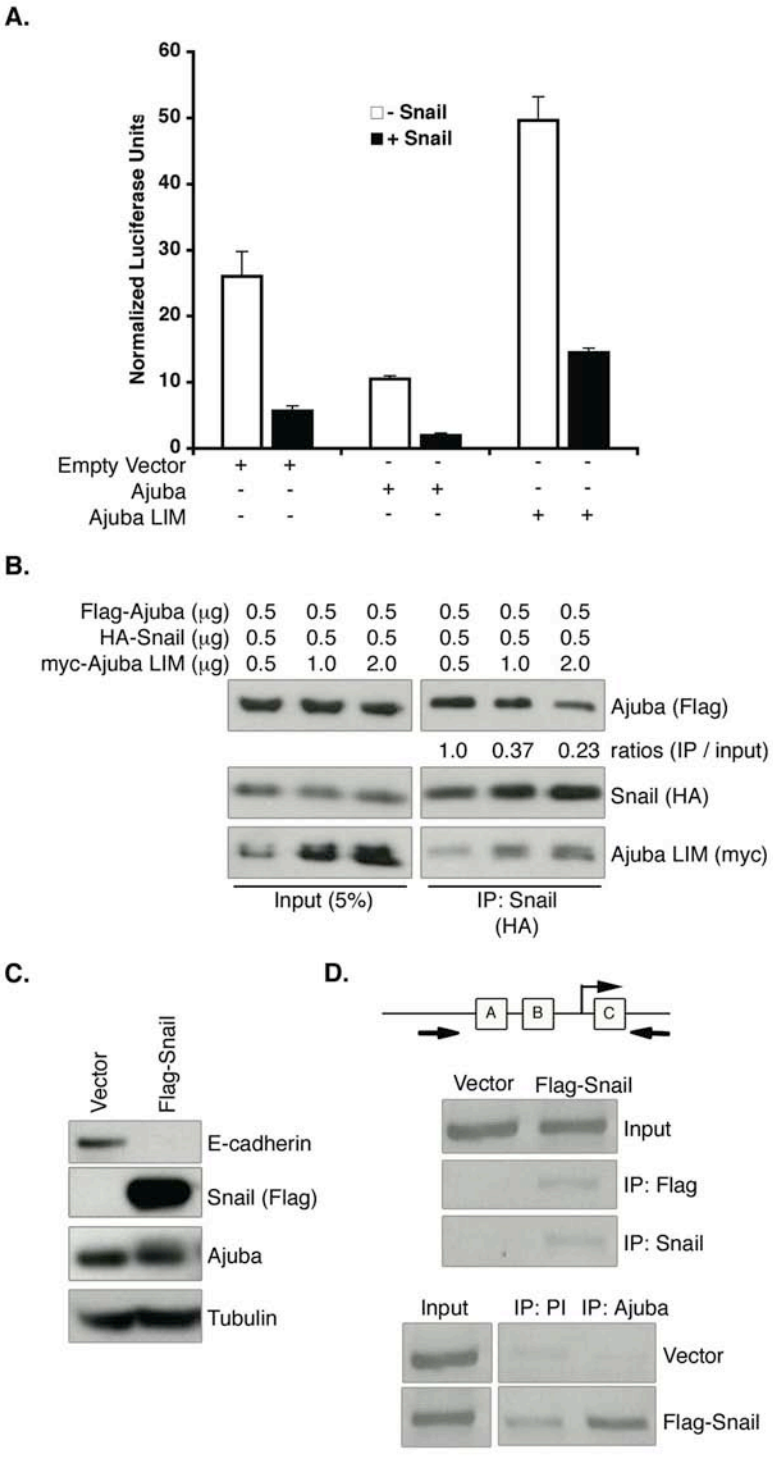
Figure 10. Model of Ajuba protein function as a Snail co-repressor. A. Ajuba interacts with Snail bound at a promoter to enhance Snail-dependent repression. The LIM region of Ajuba LIM proteins direct their interaction with, predominantly, the SNAG domain of Snail. The preLIM region is free to potentially function in recruitment of repressor complex proteins so as to allow for more efficient repression by Snail. B. In the presence of excess LIM region, endogenous full-length Ajuba LIM proteins are prevented from interacting with Snail. As a result, Snail-dependent repression is not as efficient.

Langer_Fig1

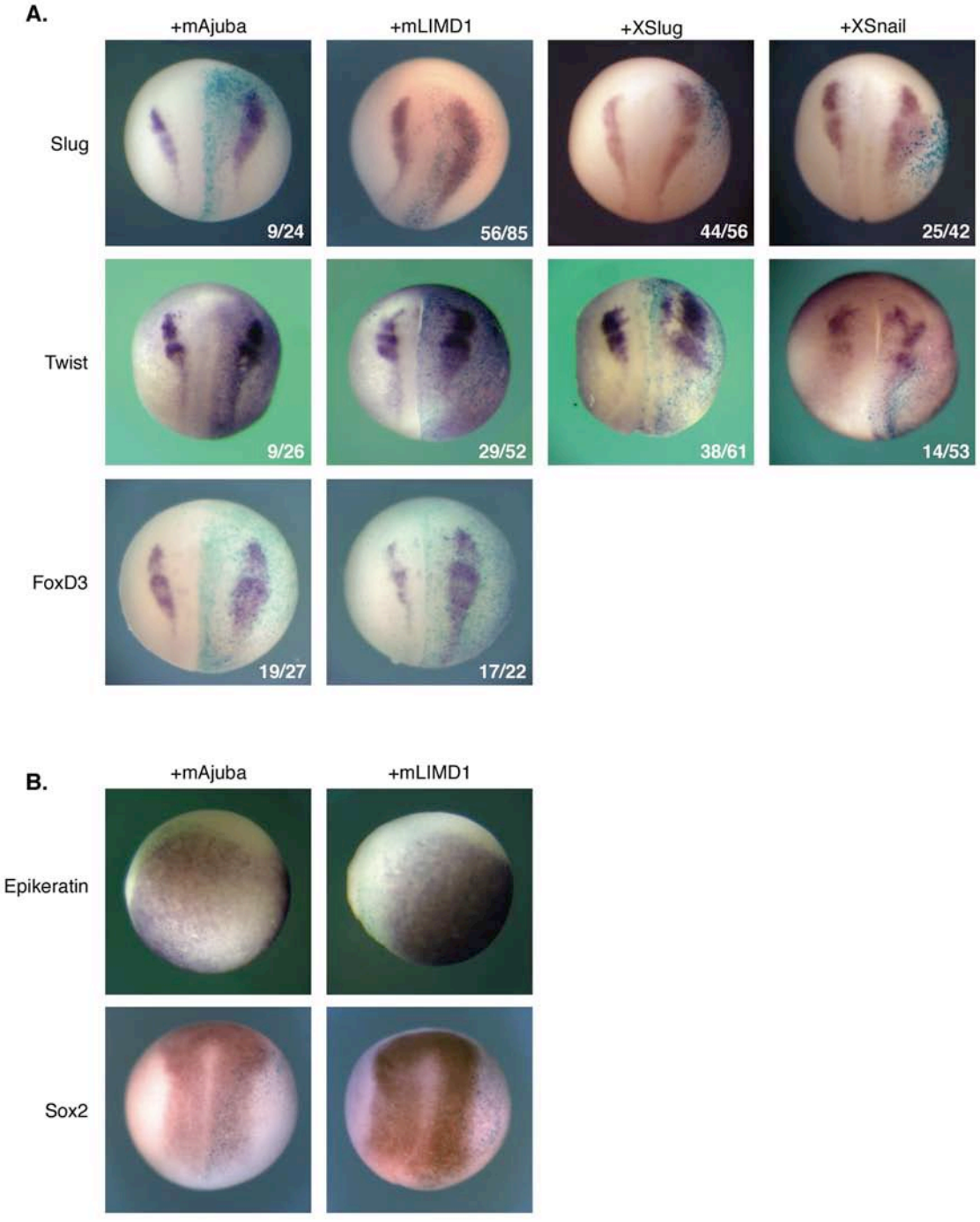

A.
myc-Ajuba + + + +
Flag-Snail - + - +
Ajuba (myc)
Snail (Flag)
IP: Snail (Flag)
Input (5%)
myc-LIMD1 + + + +
Flag-Snail - + - +
LIMD1 (myc)
Snail (Flag)
IP: Snail (Flag)
Input (5%)
myc-Zyxin + + + +
Flag-Snail - + - +
Zyxin (myc)
Snail (Flag)
IP: Snail (Flag)
Input (5%)
myc-LPP + + + +
Flag-Snail - + - +
LPP (myc)
Snail (Flag)
IP: Snail (Flag)
Input (5%)
B.
Snail Slug Scratch
Ajuba + + +
LIMD1 + + +
WTIP + ND ND
Zyxin - - -
LPP - - -
C.
NES LIM1 LIM2 LIM3
Ajuba
Ajuba PreLIM
Ajuba LIM
myc-Ajuba + + - - - - + + - - - -
myc-AjPreLIM - - + - - - - - + - - -
myc-LIMD1 - - - - + + - - - - + +
Flag-Snail - + + + - + - + + + - +
LIMD1 (myc)
Ajuba (myc)
AjPreLIM (myc)
Snail (Flag)
IP: Snail (Flag)
Input (5%)
myc-AjLIM + + + +
Flag-Snail - + - +
AjLIM (myc)
Snail (Flag)
IP: Snail (Flag)
Input (5%)
D.
SNAG
Zinc Fingers
Snail
Snail ΔSNAG
myc-Ajuba + + + + + +
HA-Snail – + – – + –
HA-ΔSNAG – – + – – +
Ajuba (myc)
1.0 0.2 ratios (IP / input)
Snail (HA)
Input (5%)
IP: Snail (HA)
E.
HaCaT MDA-231 HaCaT MDA-231 HaCaT MDA-231
LIMD1
Ajuba
Snail
Input (5%)
beads alone
IP: Snail

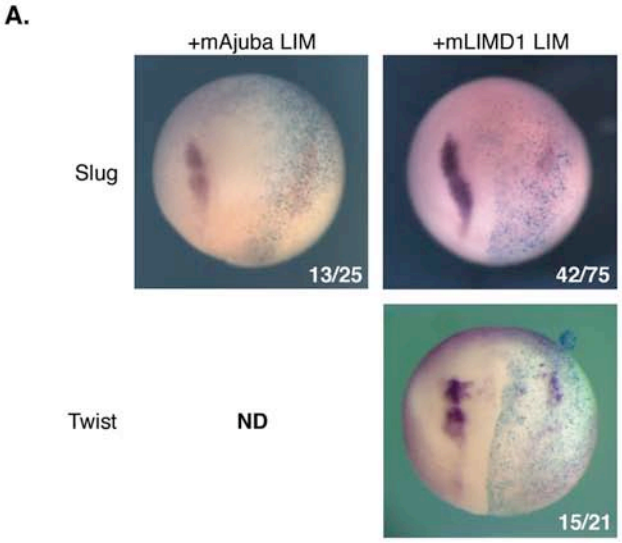
Langer_Fig2

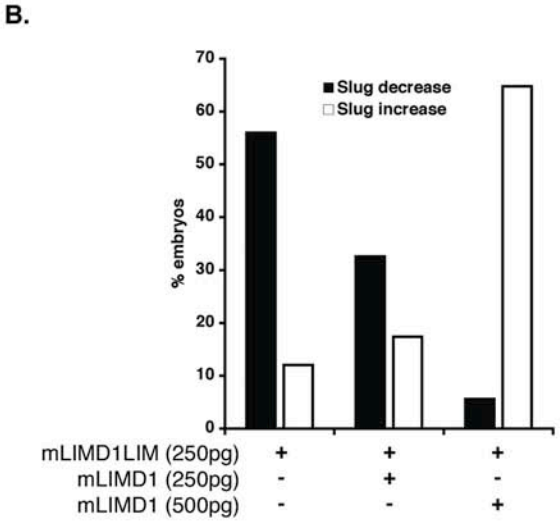

A.
RFP-Ajuba
GFP
merge
GFP
GFP-SnWT
B.
cytosolic
nuclear
% cells
100
80
60
40
20
0
GFP
GFP-SnWT
GFP-Sn8SA
C.
NE
CE
NE
CE
NE
Ajuba (myc)
Snail
Input
(2%)
IP: Snail
IP: PI


A.
Normalized Luciferase Units
60
50
40
30
20
10
0
- Snail
+ Snail
Empty Vector + + - - - -
Ajuba - - + + - -
Ajuba LIM - - - - + +
B.
Flag-Ajuba (μg) 0.5 0.5 0.5 0.5 0.5 0.5
HA-Snail (μg) 0.5 0.5 0.5 0.5 0.5 0.5
myc-Ajuba LIM (μg) 0.5 1.0 2.0 0.5 1.0 2.0
Ajuba (Flag)
1.0 0.37 0.23 ratios (IP / input)
Snail (HA)
Ajuba LIM (myc)
Input (5%)
IP: Snail (HA)
C.
Vector
Flag-Snail
E-cadherin
Snail (Flag)
Ajuba
Tubulin
D.
A
B
C
Vector
Flag-Snail
Input
IP: Flag
IP: Snail
Input
IP: PI
IP: Ajuba
Vector
Flag-Snail

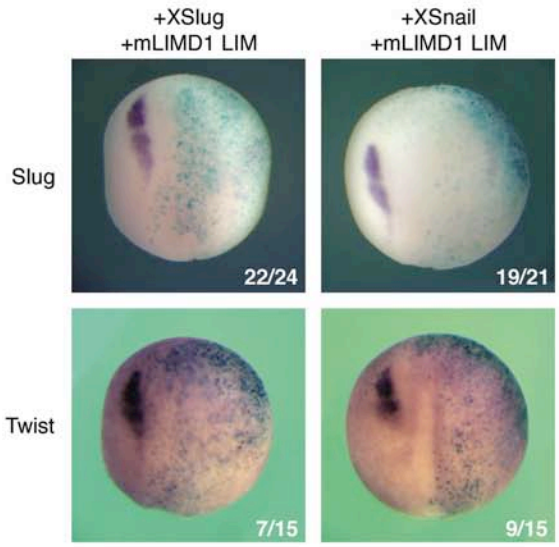
Langer_Fig4


A.
+mAjuba
+mLIMD1
+XSlug
+XSnail
Slug
9/24
56/85
44/56
25/42
Twist
9/26
29/52
38/61
14/53
FoxD3
19/27
17/22
B.
+mAjuba
+mLIMD1
Epikeratin
Sox2

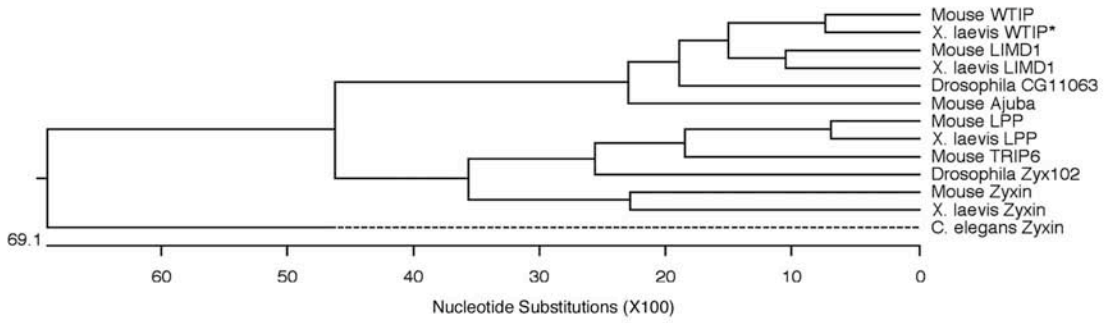

A.
+mAjuba LIM
+mLIMD1 LIM
Slug
13/25
42/75
Twist
ND
15/21


B.
Slug decrease
Slug increase
% embryos
70
60
50
40
30
20
10
0
mLIMD1LIM (250pg) + + +
mLIMD1 (250pg) - + -
mLIMD1 (500pg) - - +

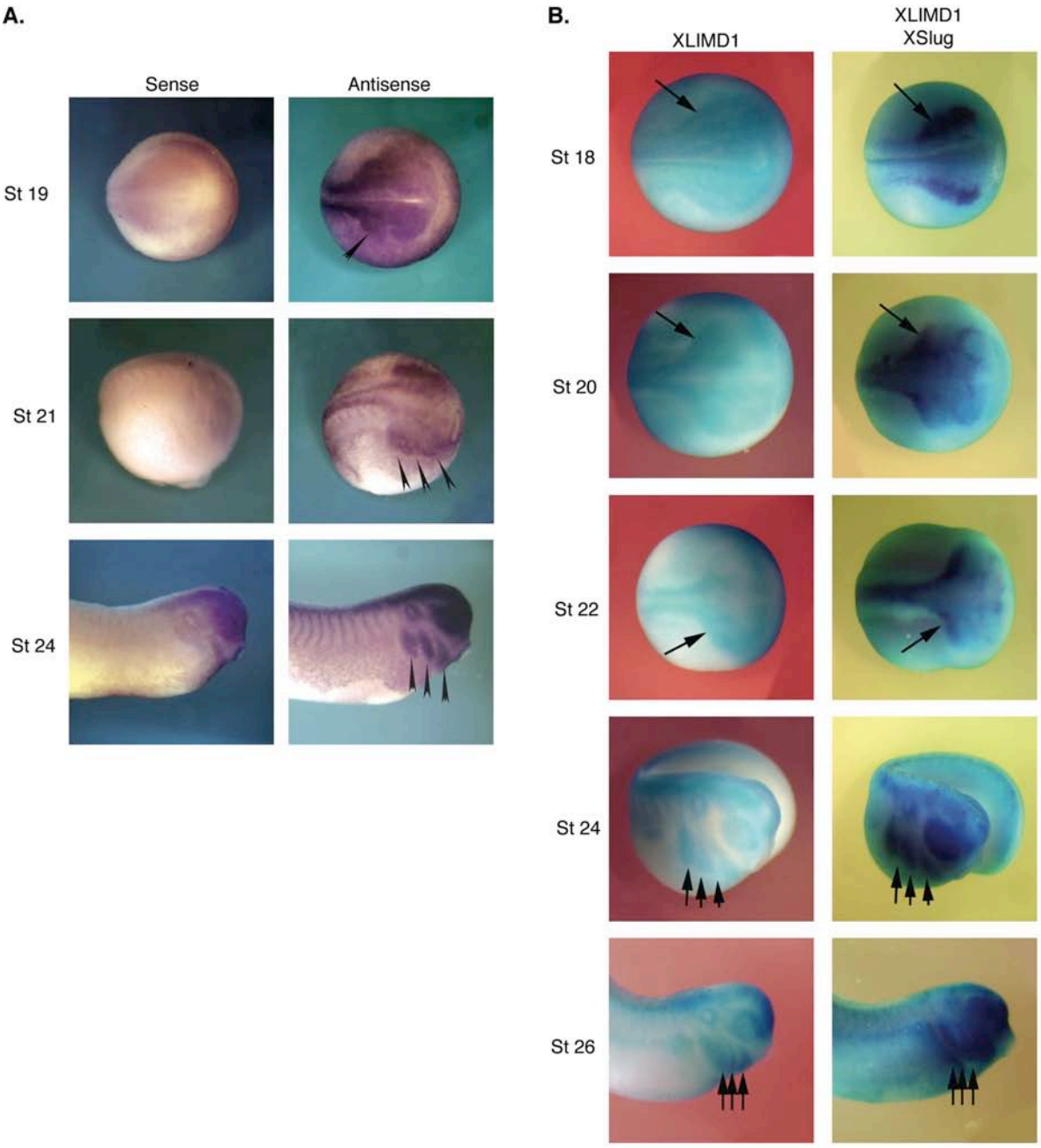
Langer_Fig6


+XSlug
+mLIMD1 LIM
+XSnail
+mLIMD1 LIM
Slug
Twist
22/24
19/21
7/15
9/15

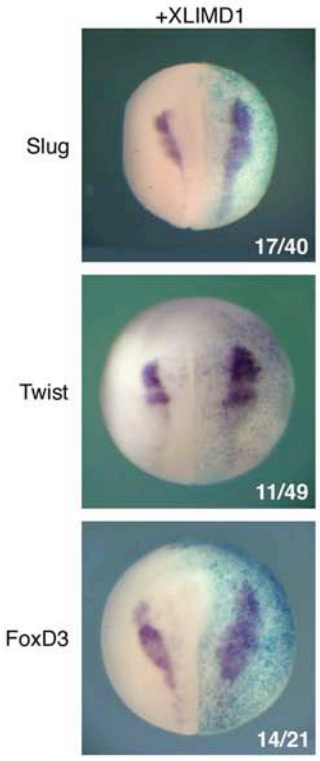
Langer_Fig7


Mouse WTIP
X. laevis WTIP*
Mouse LIMD1
X. laevis LIMD1
Drosophila CG11063
Mouse Ajuba
Mouse LPP
X. laevis LPP
Mouse TRIP6
Drosophila Zyx102
Mouse Zyxin
X. laevis Zyxin
C. elegans Zyxin
69.1
60
50
40
30
20
10
0
Nucleotide Substitutions (X100)

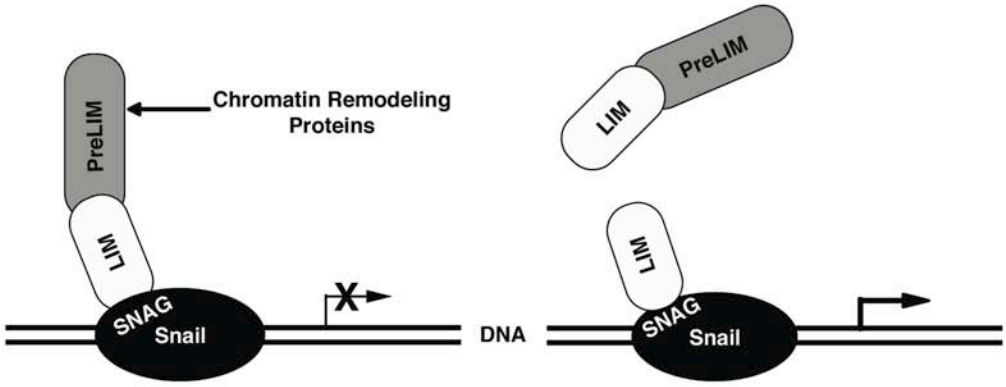
Langer_Fig8


A.
Sense
Antisense
St 19
St 21
St 24
B.
XLIMD1
XLIMD1
XSlug
St 18
St 20
St 22
St 24
St 26

Langer_Fig9



Langer_Fig10



PHD Domain-Mediated E3 Ligase Activity Directs Intramolecular Sumoylation of an Adjacent Bromodomain which is Required for Gene Silencing

Alexey V. Ivanov[1], Hongzhuang Peng[1], Vyacheslav Yurchenko[2], Kyoko Yap Kelly[3], Dmitri G. Negorev[1], David C. Schultz[1], William J. Fredericks[1], David E. White[1], Gerd G. Maul[1], Moshe J. Sadofsky[2], Ming-Ming Zhou[3], and Frank J. Rauscher, 3rd [1*].

[1]The Wistar Institute, 3601 Spruce Street, Philadelphia, Pennsylvania 19104, USA

[2]Department of Pathology, Albert Einstein College of Medicine of Yeshiva University, 1300 Morris Park Avenue, Bronx, New York 10461, USA

[3]Department of Molecular Physiology and Biophysics, Mount Sinai School of Medicine, New York University, One Gustave L. Levy Place, New York, New York 10029, USA

Running Title: PHD mediated sumoylation and gene silencing

Key Words: PHD domain; KAP1; KRAB domain; SUMO; gene silencing; bromodomain

Correspondence to:

Frank J. Rauscher, III

The Wistar Institute,

3601 Spruce Street,

Philadelphia, PA 19104

Phone: (215) 898-0995

Fax: (215) 898-3929

Email: rauscher@wistar.org

SUMMARY

Tandem PHD and bromodomains are often found in chromatin-associated proteins and have been shown to cooperate in gene silencing. While both domains can bind specifically modified histones, the mechanisms of cooperation between these domains are unknown. We show that the PHD domain of the KAP1 corepressor functions as an intramolecular E3 ligase for sumoylation of the adjacent bromodomain. The RING finger-like structure of the PHD domain is required for both Ubc9 binding and sumoylation and directs modification to specific lysine residues in the bromodomain. Sumoylation is required for KAP1 mediated gene silencing in vivo and functions by directly recruiting the SETDB1 histone methyltransferase and the CHD3/Mi2 alpha component of the NuRD complex via highly conserved SUMO interacting motifs (SIMs). Furthermore, sumoylated KAP1 stimulates the histone methyltransferase activity of SETDB1. Our data provide a novel mechanistic explanation for cooperation of PHD and bromodomains in gene regulation and describe a new function of the PHD domain as an E3 SUMO ligase.

INTRODUCTION

Eukaryotic gene regulation requires the coordination of multiple enzymatic activities that target nucleosomal histones for post-translational modification. Protein domains which recognize specific histone modifications are key components of this mechanism (Seet et al., 2006). The bromodomain recognizes acetylated lysines in histone H3 and H4 (Yang, 2004). The PHD domain of ING2 and BPTF binds methylated histone H3 (Li et al., 2006; Shi et al., 2006; Wysocka et al., 2006). The PHD and bromodomain are often found adjacent to each other suggesting that they may form a cooperative module used to interpret combinatorial modifications of histones. Cooperation of the PHD and the bromodomain occurs in p300, TIP5, BPTF and KAP1 proteins (Ragvin et al., 2004; Schultz et al., 2001; Wysocka et al., 2006; Zhou and Grummt, 2005).

KAP1/TIF1 beta is a universal corepressor for the largest family of transcriptional silencers encoded in the human genome, KRAB domain containing zinc finger proteins (ZFPs) (Huntley et al., 2006). The KAP1 N-terminal RBCC region is responsible for KRAB domain binding (Friedman et al., 1996; Peng et al., 2000), while HP1 binding domain and PHD-Bromodomain tandem are required for gene silencing. The KAP1 PHD-Bromodomain is required for recruitment of the H3-K9 specific HMTase SETDB1 and CHD3 proteins to the promoter regions of KRAB modulated genes (Schultz et al., 2002; Schultz et al., 2001; Sripathy et al., 2006). Mutations in either the PHD or bromodomains compromise KAP1 interaction with these partner proteins and relieve repression, suggesting that they act cooperatively. The PHD domain of KAP1 is highly related to the RING finger (Capili et al., 2001), often found in ubiquitin and SUMO E3 ligases (Joazeiro and Weissman, 2000).

SUMO family proteins are conjugated to target lysines via a cascade of E1 activating, E2 transfer and E3 ligase enzymes (Johnson, 2004). The SUMO E2 protein Ubc9 often recognizes the consensus sequence ψKxE/D (ψ - hydrophobic) in the target protein and catalyzes SUMO conjugation (Sampson et al., 2001) leading to alterations in protein localization and function. A family of deconjugation enzymes, SENP1-7, is responsible for rapid removal of SUMO from target lysines (Mukhopadhyay and Dasso, 2007) which accounts for the very temporal nature of this modification. Sumoylation is emerging as a key posttranslational modification involved in transcriptional repression (Gill, 2005), however, the detailed mechanisms of its action remain obscure. SUMO modification can enhance recruitment of repression machinery, including HDACs (Girdwood et al., 2003; Yang and Sharrocks, 2004) and may also act to preclude other lysine modifications leading to impaired protein-protein interactions.

Here, we show that the KAP1 PHD domain binds to Ubc9 and directs SUMO conjugation of an adjacent bromodomain, and that sumoylation is required for KRAB-KAP1 mediated repression. The tandem PHD-Bromodomain of KAP1 functions as an intramolecular SUMO E3 ligase. SETDB1 and CHD3 encode functional SIM motifs and only bind SUMO-modified KAP1. Thus, sumoylation mediated repression occurs via direct recruitment of H3-K9 HMTase and HDAC activities.

RESULTS

KAP1 is Sumoylated In Vivo and In Vitro

SUMO is often conjugated to lysine residues within the consensus ψKxE/D. The KAP1 sequence contains six lysines conforming to this consensus: canonical sites at K554 and K676, and four reverse sites at K575, K750, K779 and K804 (Figure 1A,B). Immunoprecipitation of endogenous KAP1 with two different antibodies under conditions preserving the SUMO modification revealed a ladder of SUMO1-modified KAP1 species (Figure 1C). Expression of SUMO1 and Ubc9 or a dominant negative SENP1 greatly enhanced KAP1-SUMO1 conjugation, whereas expression of wild type SENP1 prevented it (Figure 1D). KAP1 sumoylation was also efficiently reconstituted in vitro with purified proteins. Reaction containing all components produced a ladder of slower migrating KAP1 forms consistent with conjugation of multiple SUMO1 moieties. Reactions that lacked either E1 or Ubc9 did not show sumoylation products (Figure 1E). Thus, KAP1 is conjugated with SUMO on multiple sites, both in vivo and in vitro.

Mutation of the PHD Domain Inhibits KAP1 Sumoylation

The use of truncated proteins showed that SUMO modification occurs only in the C-terminal half of KAP1 (NPB fragment, see Figure 1F,G) overlapping the known repression domains: NHD (**N**-CoR2 **H**omology **D**omain, see Figure S1A), the PHD and the bromodomain. A structural analysis of the KAP1 PHD domain showed it to be highly related to the RING finger (Capili et al., 2001), often found in ubiquitin ligases and in SUMO E3 ligases of PIAS family

(Joazeiro and Weissman, 2000). The RING domain in PIAS proteins is responsible for interaction with Ubc9 and its structural mutation leads to inhibition of sumoylation activity ((Kotaja et al., 2002) and Figure 3F). To test the hypothesis that the PHD domain can function in a similar way to the RING finger of the PIAS ligases, we introduced a C651A mutation, which disrupts its structure (Capili et al., 2001). SUMO modification of the C651A mutant protein was drastically reduced (Figure 1F,G, lane 4). However, the intact PHD was required for sumoylation of many, but not all sites in KAP1 (Figure 1F,G, arrow). Sumoylation of the NP fragment which lacks the bromodomain was unaffected by C651A mutation (Figure 1H). Therefore, mutation in the PHD affected sumoylation of the bromodomain, but not the adjacent NHD domain or the PHD itself. These results suggest that the PHD domain plays a key role in KAP1 sumoylation and it is required to target the adjacent bromodomain.

Selective Utilization of Sumoylation Sites and Their Contribution to KAP1 Repression Activity

First, we confirmed putative KAP1 sumoylation sites by lysine-to-arginine mutational analysis and assessed the contribution of sumoylation to the repression function of KAP1. In agreement with our previous report (Schultz et al., 2001), the NHD, PHD and bromodomain all demonstrated significant repression. Mutation of the canonical consensus lysine K554 in the NHD resulted in reduction in sumoylation and repression (Figure 2A,B, lane 2). Mutation of K575 modestly reduced, while the double mutation abolished, both sumoylation and repression (lane 4). Mutation of K676 in the PHD domain had a similar effect (lane 6). Simultaneous mutation of K779 and K804 in the bromodomain completely abrogated sumoylation and

repression (lanes 8-10). Introduction of triple substitutions in the PB (**P**HD-**B**romodomain) module revealed an additional, cryptic acceptor lysine K750 in the bromodomain which was not active in repression assay in the absence of the PHD (lane 10) and occurs in a non-canonical motif, QE**K**L. Additional inactivation of this fourth site completely abrogated repression (Figure 2C, lane 7). It is noteworthy that mutation of other lysines in the bromodomain, K770 and K774 had no such effect (lane 8). Analysis of mutant PB proteins with only one wild type acceptor lysine (out of four) allowed us to evaluate relative contribution of each site to sumoylation efficiency and repression activity (Figure 2C,D). Lysines K779 and K804 were predominant conjugation sites and most efficiently recapitulated repression of the wild type protein. Although SUMO conjugation to K676 and K750 was much less efficient we confirmed that sumoylation of these sites contributed to repression. We took advantage of a specific inhibitor of the endogenous sumoylation machinery, Gam1, which targets the SUMO E1 for degradation (Boggio et al., 2004). When each mutant was cotransfected with the wild type Gam1, the repression activity was strongly reduced (Figure 2E), suggesting that each lysine is a site for in vivo sumoylation and that, irrespective of other modifications which may occur on these lysines, sumoylation has a major contribution to repression.

Similar analysis of the NPB fragment of KAP1 containing all six acceptor lysines allowed us to establish a strict hierarchy in the site utilization. The most efficiently sumoylated sites were K779, K804 and K554, and less actively utilized - K575, K676 and K750 (Figure 2F,G). Moreover, the difference in SDS-PAGE mobility of SUMO-conjugated NPB mutants helped us to assign specific bands to a particular acceptor lysine modification. Thus, sumoylation of the wild type NPB fragment revealed one major mono-sumoylation product and several minor slower migrating bands (lane 1). This predominant, mono-sumoylated NPB protein migrated in

identical position to the SUMO conjugates at the bromodomain sites K779 or K804 (lanes 6,7; see also Figure S1B), suggesting that sumoylation of KAP1 primarily occurs on these two sites. Again, use of mutant NPB proteins with only one wild type acceptor lysine (out of six) showed a very good correlation between the extent of sumoylation and repression activity (Figure 2F-H). Mutation of all six acceptor lysines almost completely abolished sumoylation and repression (lane 8). Thus, sumoylation of KAP1 is required for its repression function and plays pivotal role in KAP1 mediated silencing.

The KAP1 PHD Binds to Ubc9 with High Selectivity and Directs Sumoylation of the Adjacent Bromodomain

Next, we investigated whether the PHD domain is indeed a SUMO E3 ligase. The PHD-Bromodomain module (PB) was a >10-fold better substrate for sumoylation and better repressor than either domain alone (Figure 2A,B, compare lanes 5,7 and 12). Mutation of the PHD (C651A) greatly diminished SUMO conjugation to the bromodomain sites K779 and K804 and impaired repression (Figure 3A,B, lanes 5-7) implying that PHD enhances bromodomain sumoylation.

To determine if the PHD domain can directly associate with the SUMO E2 enzyme, Ubc9, or other components of the SUMO conjugation machinery we performed GST binding experiments with purified proteins. Recombinant PHD protein bound strongly and stochiometrically to Ubc9 but did not bind to SUMO1, Sae1 or Sae2 (Figure 3C). Structural mutations in the PHD domain, C651A or W664A completely abolished binding to Ubc9 (Figure 3D, lanes 3,6,9). Interestingly, the PHD domains from the closely related TIF1/KAP1 family

members TIF1α and TIF1γ as well as from the chromatin protein Sp100C were unable to bind to Ubc9 (Figure S2A). Moreover, the PHD-Bromodomain (PB) modules from TIF1α, TIF1γ and Sp100C were very poor sumoylation substrates compared to the KAP1 PB (Figure 3E) despite the fact that they contain similar sumoylation consensus sites (Figure S2B). Thus, the correlation between Ubc9 binding and sumoylation provides strong evidence that KAP1 PHD-Ubc9 interaction is critical for KAP1 SUMO modification.

To compare KAP1 with the known SUMO E3 ligase PIAS1 we analyzed their sumoylation efficiency in relation to the structural integrity of the Ubc9 binding determinants, PHD and RING domains, respectively. Both, RING and PHD domains were poor substrates themselves (Figure 3F, lanes 1 and 5). However, they efficiently stimulated sumoylation of adjacent domains, N-terminal globular domain in PIAS1 (RR) and bromodomain in KAP1 (PB), respectively (lanes 3 and 7). Structural mutations in both RING and PHD domains, which disrupt Ubc9 binding, drastically reduced PIAS1 and KAP1 sumoylation (lanes 4,8) suggesting that the KAP1 PHD domain functions similarly to the PIAS1 RING domain and acts as an intramolecular SUMO E3 ligase.

To distinguish whether the KAP1 PHD can function as a SUMO ligase in-trans as well as in-cis, we tested its ability to stimulate sumoylation of the bromodomain in reactions containing the two domains as separated, individual molecules. Addition of increasing amounts of PHD protein to the bromodomain or PB with mutated PHD domain did not result in increased SUMO conjugation (Figure S2C). However, when PHD was fused to the C-terminus of the bromodomain (Bromo-PHD chimera) sumoylation was significantly enhanced (Figure 3G, compare lanes 2 and 4) indicating that covalent link between the two domains is required for proper SUMO ligation.

Mapping of Ubc9 Binding Site on the KAP1 PHD Domain

To map interaction interface in the KAP1 PHD involved in Ubc9 binding we performed NMR titration experiments. As shown in the 2D ^{1}H-^{15}N HSQC (heteronuclear single quantum correlation) NMR spectra of the ^{15}N-labeled KAP1 PHD domain (Figure 4A), many amino acid residues of the protein exhibited significant chemical shift perturbations - predominantly line-broadening effect - upon addition of Ubc9 in a concentration dependent manner, indicative of direct PHD-Ubc9 interactions. The ligand binding induced line-broadening effect likely results from the dissociation rate of the PHD-Ubc9 complex on the millisecond NMR timescale, suggesting a low micromolar binding affinity of the two proteins. By the NMR resonance assignments (Figure S3) we further mapped the Ubc9 binding site on the protein and color-coded the most perturbed amino acid residues on the surface representation model of the PHD molecule (Figure 4B). An exhaustive mutational analysis of the KAP1 PHD domain sequence, which included almost every amino acid substitution, revealed a very good correlation between decreased in vitro sumoylation efficiency of a particular missense mutant and the NMR resonance perturbation of corresponding residue upon Ubc9 titration (Figure S4 and 4C).

Interestingly, most of the Ubc9 perturbed residues are located largely on one side of the PHD molecule and belong to both zinc coordination sites, as well as to loop 1 between β_1 and β_2 strands, loop 2 between β_3 and β_4 strands and to an adjacent region containing H652 and L653 (Figure 4C) at the PHD/bromodomain interface (see accompanying paper, Zeng et al). The loop 1 perturbed residues are well conserved among PHD domains of TIF1 family members, whereas the loop 2 and β_4 strand residues, and L653 are unique for the KAP1 PHD (Figure 4D and S5A).

We performed subdomain swaps of the most diverged contiguous amino acid sequences between the KAP1 and TIF1γ PHD domains where KAP1 PHD segments containing loops 1, 2 or 3 were replaced with corresponding TIF1γ sequences (6 aa substitutions) (Figure 4D). The respective mutants, L1, L2 or L3 were tested in GST pull-down, in vitro sumoylation and repression assays (Figures 4E-G) which confirmed that the L2 segment of the KAP1 PHD together with L653 contain the critical amino acid contacts/surfaces required for binding to Ubc9. We verified that the L653A and L2 substitutions did not cause any gross changes in the PHD structural fold that were observed for the C651A mutant (Figure S5B). Together these results provide strong evidence that the KAP1 PHD directly binds to Ubc9 and can function as a SUMO E3 ligase.

Sumoylation Is Required for KRAB Domain Mediated Repression

Since KAP1 is an obligate corepressor for the KRAB domain, we first sought to determine if sumoylation is required for KRAB domain mediated repression. For a comparison, along with the KRAB several other well characterized repression domains were tested. LexA fusions to KRAB, POZ, SNAG and FOG domains were cotransfected with the Gam1, a specific inhibitor of the endogenous sumoylation pathway (Boggio et al., 2004). Each domain demonstrated potent repression in the absence of Gam1 (Figure 5A). However, the repression activity of LexA-KAP1 PB and LexA-KRAB was markedly reduced by wild type Gam1 but not the Gam1 mutant. The effect of Gam1 on repression was specific for the KRAB-KAP1 pathway since neither POZ, SNAG nor FOG domains showed significant relief of repression in the presence of Gam1. We confirmed that expression of Gam1 markedly decreased KAP1 sumoylation in cells (Figure 5B). This observation emphasizes the importance of sumoylation for

KRAB domain repression function and sets it apart from mechanisms of action of the other repressor domains.

Sumoylated KAP1 is Required to Complement KRAB-mediated Repression in KAP1 Knockdown Cells

To verify that sumoylation of KAP1 in particular is required for KRAB mediated repression we created U2OS cell lines with a stable knockdown of KAP1 using vector-based shRNA technology (Wang et al., 2005). Clone U2OS-K4 showed a >95% knockdown of endogenous KAP1 (Figure S6A and 5C). Cotransfection of a GAL4-TK-luciferase and Gal4-DBD or Gal4-KRAB plasmids into control U2OS cells yielded about 10-fold repression, whereas this repression activity in U2OS-K4 cells was significantly attenuated (Figure 5C). Complementation with a FLAG-KAP1 cDNA resistant to the action of the shRNA (Figure S6B) restored Gal4-KRAB mediated repression to 18-fold. In contrast, the sumoylation deficient mutant of KAP1, K6R, which has each of the six SUMO acceptor lysines mutated to arginines, had no stimulatory effect on Gal4-KRAB repression activity (Figure 5D). Both wild type KAP1 and the mutant K6R protein were properly localized to pericentromeric heterochromatin (Figure S6C), indicating that the loss of repression was not due to defect in the K6R subnuclear targeting. The KAP1 PHD mutant, C651A which showed an intermediate level of sumoylation (Figure 1F,G) demonstrated poor ability to complement Gal4-KRAB-mediated repression. A fusion of the K6R protein to SUMO1 had a very modest ability to complement KRAB mediated repression indicating that multiple sumoylation at naturally occurring sites in KAP1 is required

for full activity. Altogether, these results strongly imply that KAP1 sumoylation is essential for its corepressor function.

SETDB1 and CHD3 Interact Only with Sumoylated KAP1 in Yeast

Repression mediated by the KAP1 PHD-Bromodomain (PB) module depends on its interaction with the H3-K9 histone methyltransferase SETDB1 and a component of histone deacetylase complex - chromatin remodeling factor CHD3 (Schultz et al., 2001; Sripathy et al., 2006). We asked whether these interactions are influenced by sumoylation of KAP1. Yeast two hybrid assay showed robust interactions of wild type LexA-KAP1 PB with Gal4 activation domain fusions to SETDB1 or CHD3 as indicated by yeast growth under the appropriate selective conditions (Figure 6A). However, LexA-PB with the K676R mutation interacted with CHD3, but not with SETDB1, suggesting that sumoylation of K779 site in yeast is not efficient enough to score SETDB1 binding. Simultaneous mutation of K676 and K779 sites abolished interaction with both SETDB1 and CHD3. Immunoprecipitation and Western blot analysis of the KAP1 bait proteins showed that only K676 and to a much lesser extent K779 sites were subject to sumoylation in yeast (Figure 6B, lanes 1-4). Unlike in mammalian cells, the yeast SUMO protein, Smt3p formed poly-Smt3p chains on KAP1, consistent with published reports (). When the LexA-KAP1 NPB baits were analyzed, we observed that only lysines K554 and K676 were required for association with SETDB1 and CHD3 (Figure 6A). Among the six identified sumoylation sites in KAP1, only K554 and K676 conform to the canonical consensus (Figure 1A) suggesting that sumoylation in yeast preferentially occurs on canonical sites. These results strongly indicate that sumoylation of KAP1 is required for binding to SETDB1 and CHD3.

Identification of Functional SIM Motifs in SETDB1 and CHD3

To determine if binding of SETDB1 or CHD3 to SUMO may be sufficient to score genetic interaction with sumoylated KAP1, we used LexA fusions to human SUMO1, SUMO2 and the SUMO1mt (K37A,K39A). SETDB1 and CHD3 interacted with SUMO1 and SUMO2, but not with the SUMO1mt (Figure 6D). This mutant is a well-characterized, two amino acid alteration of the binding pocket of the SUMO molecule that abolishes its interaction with SIM (SUMO Interacting Motif) found in proteins that bind SUMO. We hypothesized that SETDB1 and CHD3 each contain SIM motif and used extensive mutational analysis to identify it. A core SIM with the consensus sequence V/I-X-V/I-V/I, (X - any amino acid) has been identified and structurally characterized (Chupreta et al., 2005; Hecker et al., 2006; Song et al., 2005). Deletion of 23 aa from the N-terminus of SETDB1 fragment abolished yeast growth and reporter gene expression (Figure 6C). The sequence of this 23 aa peptide encodes a 4 aa segment (122-IIEI-125) conforming to the core SIM consensus (Figure 6E). When this putative SIM was mutated the interaction with KAP1 PB was abolished. A 35 aa C-terminal truncation of CHD3 abolished its interaction with KAP1 PB (Figure 6C). This sequence encodes a 4 aa segment (1995-VICI-1998) conforming to the core SIM consensus. When the six C-terminal residues of CHD3 containing this putative SIM were deleted the interaction with KAP1 PB was completely abrogated. These experiments identified bona fide SIM domains in SETDB1 and CHD3. In both proteins, the SIM domains have been highly conserved throughout vertebrate evolution (Figure S7A) suggesting that they play an important functional role.

SETDB1 and CHD3 Bind to Sumoylated KAP1 In Vitro in a SIM Dependent Manner

To confirm that these interactions are direct, we reconstituted them in vitro. *E.coli* derived GST fusion proteins containing a high proportion (>50%) of properly sumoylated forms were used in binding assays (Figure 6F, bottom, brackets). SETDB1 showed moderate binding to SUMO1 and this association was abolished by the K37A,K39A mutation in SUMO1 (SUMO1mt) (lane 4 and Figure S7B). SETDB1 bound poorly to unmodified PB but the interaction was dramatically stimulated when sumoylated PB (PB-S1) or sumoylated full length KAP1 (KAP1-S1) were used (lanes 6,8). SETDB1 which lacks the SIM motif (SETDB1dSIM) showed markedly reduced binding to all sumoylated proteins (middle). SETDB1 showed very little affinity to highly sumoylated Sp100A (Sp100-S1,lane 10) suggesting that the SUMO modification alone is not the sole determinant for interaction in all cases.

We observed similar results in reverse binding experiments (Figure 6G). Unmodified KAP1 did not bind to any of the GST-SETDB1 or GST-CHD3 resins (lanes 2-7). In contrast, sumoylated KAP1 was efficiently retained by wild type SETDB1 and CHD3, while mutation of their SIM motifs completely abolished binding (lanes 10-13). Interestingly, multi-sumoylated KAP1 conjugates (higher molecular weight bands) bound to SETDB1 and CHD3 more efficiently than mono-sumoylated conjugates despite their lower abundance in the preparation, providing biochemical evidence to explain our in vivo data that multi-sumoylation of KAP1 is required for efficient KRAB-KAP1 mediated repression. Together, these data indicate that SUMO-SIM interaction provides a major contribution to the interaction between sumoylated KAP1 and SETDB1 or CHD3.

Endogenous KAP1 binds to SETDB1 and CHD3 in a Sumoylation Dependent Manner

To assess the role of sumoylation for KAP1-SETDB1/CHD3 interaction in vivo we reintroduced wild type or different functional KAP1 mutants into our U2OS-K4 KAP1 knockdown cells. Specifically, we selected clonal cell lines expressing exogenous FLAG-KAP1 proteins at nearly physiological levels (Figure 7A, compare lanes 1 and 4-8). The reconstituted KAP1 mutants included C2 (abolishes trimerization/KRAB binding) (Peng et al., 2000), M2 (abolishes HP1 binding/subnuclear targeting) (Ryan et al., 1999), the K6R (abolishes sumoylation) and dB (deletion of the bromodomain, decreases sumoylation). Cotransfection of a GAL4-TK-luciferase and Gal4-DBD or Gal4-KRAB plasmids into K4/WT cells yielded about 17-fold repression, whereas this repression activity in control K4/vector and in K4/C2, K4/M2, K4/K6R and K4/dB cells was significantly attenuated (top). This is consistent with our previous data showing that KAP1 trimerization (abolished by the C2 mutation), HP1 binding/subnuclear targeting (abolished by the M2 mutation) and sumoylation (abolished by the K6R mutation, decreased by deletion of the bromodomain (dB)) (Figure 5D) are critical for KAP1 repression function. Surprisingly, in addition to the bona fide sumoylation deficient mutant, K6R all the other KAP1 functional mutants showed dramatically reduced levels of SUMO modification in vivo by both, SUMO1 and SUMO2 (Figure 7B, middle, lanes 14, 16-18) suggesting that impairment in sumoylation might be the major factor in the loss of their repression activity. We detected efficient co-immunoprecipitation of endogenous SETDB1 and Mi2/CHD3 with wild type KAP1 only when cells were lysed and processed in conditions preserving the SUMO modification (+NEM) (top, compare lanes 15 and 21). NEM is an inhibitor of de-sumoylating SENP enzymes, which deconjugate SUMO from substrate proteins upon cell lysis (lanes 7-12).

Consistent with this, there was no or very little association between KAP1 and Mi2/CHD3 or SETDB1 in the absence of NEM (lanes 20-24). Moreover, markedly reduced sumoylation of the KAP1 mutants resulted in lower efficiency of SETDB1 and Mi2/CHD3 co-immunoprecipitation (lanes 14,16-18). Together these data suggest that KAP1 sumoylation plays a major role in its interaction with SETDB1 and CHD3 in vivo.

KAP1 Sumoylation is Required for Recruitment of SETDB1 to the Target Promoter

To confirm this conclusion we tested the requirement of KAP1 sumoylation for the recruitment of SETDB1 to target promoter. We introduced a GAL4-TK-luciferase, Gal4-DBD or Gal4-KRAB and various KAP1 variants into U2OS-K4 KAP1 knockdown cells and performed chromatin immunoprecipitation (ChIP) assay with primers specific to the TK promoter. As expected, all Gal4 fusion proteins bound to the promoter (Figure 7C, panel a). Consistently, the wild type (WT), K6R and M2 mutant KAP1 proteins were efficiently recruited to the promoter by Gal4-KRAB but not by Gal4-DBD (panel b, lanes 5,7,9). However, SETDB1 and the histone H3 K9-trimethyl mark were only enriched at the promoter in cells transfected with wild type KAP1 or control vector (panels c,d, lanes 3,5). Interestingly, since HP1 binding to KAP1 does not directly depend on KAP1 sumoylation (Ryan et al., 1999), HP1α was recruited to the promoter by the K6R mutant as efficiently as by wild type KAP1, but it did not result in appearance of the histone H3 K9-trimethyl mark and subsequent repression. Conversely, the HP1-binding deficient mutant, M2 did not recruit HP1 and neither SETDB1 nor the histone H3 K9-trimethyl mark (panels c-e, lane 9), since it is also defective in sumoylation (Figure 7B, lane

17). Together, these results suggest that KAP1 sumoylation is required for the SETDB1 recruitment to the target promoter and the establishment of repressive chromatin marks.

Sumoylated KAP1 Stimulates SETDB1 HMTase Activity in Vitro

To investigate whether sumoylated KAP1 influences SETDB1 enzymatic activity, we performed in vitro histone methyltransferase (HMT) assay in the presence of unmodified or sumoylated KAP1 PB module (see Figure S8A). Low levels (100 ng) of baculovirus-expressed SETDB1 were incubated with core histones in a standard HMTase reaction containing increasing amounts of KAP1 PB. Under these conditions, SETDB1-mediated methylation of histone H3 was undetectable in the absence of KAP1 (Figure 7D, lane 2). (A twenty fold higher amount of SETDB1 (2 μg) was required to produced robust methylation of the histone H3 (data not shown)). When unmodified KAP1 PB was added to the reaction, a low level of H3 methylation was detected (lane 5). However, the addition of sumoylated KAP1 PB resulted in a robust, dose-dependent increase in histone H3 methylation (lanes 6-8), suggesting that sumoylated KAP1, which binds SETDB1, enhanced its catalytic activity. Similar titration reactions with SUMO1 protein alone did not result in stimulation of SETDB1 (Figure S8B) suggesting that KAP1 and SUMO sequences together contribute to SETDB1 activation. In summary, these data strongly indicate that sumoylated KAP1 can increase SETDB1 HMTase activity toward histone H3.

DISCUSSION

KAP1 PHD is an Intramolecular SUMO E3 Ligase for the Adjacent Bromodomain

Tandem PHD and bromodomains are often found in chromatin associated proteins and have been shown to cooperate in gene regulation. Based upon recent structural analyses, it has been proposed that tandem PHD and bromodomains can independently recognize modifications in histones, and thus may act combinatorially to interpret the histone code (Li et al., 2006). In this article we provide an alternative basis for the cooperativity of these domains in corepressor protein KAP1 by demonstrating that the PHD domain can function as a SUMO E3 ligase in-cis for the bromodomain.

We found that the PHD domain of KAP1 functions in a similar way to the RING finger of the PIAS SUMO E3 ligases. It directly binds to the SUMO E2 enzyme Ubc9 and is necessary for KAP1 sumoylation. The immediate target of the KAP1 PHD ligase activity is the adjacent bromodomain. In addition to binding to the E2 enzyme a true E3 ligase has to fulfill the second criteria – interact with its target. The solution structure of the KAP1 PHD-Bromodomain tandem confirmed that these two domains physically interact and form a structural and functional module (Zeng et al., see accompanying paper). This interaction seems to be important since chimeras between KAP1 PHD and bromodomains from other TIF1 family members fail to fully recapitulate repression activity (Schultz et al., 2001).

It seems likely that the KAP1 PHD may also function as a SUMO E3 not only intramolecularly but also in-trans, by enhancing sumoylation of other KAP1 interacting proteins. Interestingly, we have not detected KAP1 dependent stimulation of SETDB1 and CHD3 protein

sumoylation implying that they are effector molecules for KAP1 but are not direct targets for the -E2-E3 (Ubc9-PHD) cascade.

Sumoylation Dependent Interaction of KAP1 with Its Effector Proteins

It has been demonstrated that controlled recruitment of a Gal4-fusion of KAP1 to a chromatinized GAL4-TK-luciferase transgene results in targeted localization of SETDB1 to the promoter region, a local increase in histone H3-K9 tri-methylation and repression. Transient knockdown of SETDB1 in these conditions compromises repression and thus provides evidence for functional link between KAP1 repression activity and SETDB1 recruitment (Sripathy et al., 2006). Here we provide direct evidence that the interaction between KAP1 and SETDB1 or CHD3 depends on KAP1 sumoylation and is a function of their specific SIM motifs. We show that SETDB1 and CHD3 are able to interact with SUMO1 and SUMO2. Consistent with this finding, SETDB1 was recently identified among other bound proteins in GST-SUMO2 affinity chromatography purifications (Rosendorff et al., 2006). It appears that specific recognition of SUMO-conjugated transcriptional regulators by individual components of repression machinery could be a general phenomenon. It has recently been shown that another partner protein of SETDB1, MCAF1, contains SIM domain which facilitates its association with sumoylated MBD1 (Uchimura et al., 2006). Similarly, recruitment of the DAXX protein by SUMO-modified glucocorticoid receptor is dependent on its SIM domain (Lin et al., 2006).

Our discoveries are summarized in a schematic model describing the mechanism of KRAB-KAP1 mediated repression (Figure 7D). Like other SUMO substrates, it appears that only a small percentage of total cellular KAP1 may be SUMO-modified at any one time. We favor the

hypothesis that sumoylation of KAP1 occurs in a controlled, transient fashion only at specific sites in chromatin, for example upon KRAB-ZFP-KAP1 interaction at the target gene promoters. The multiple KAP1-SUMO species observed in vivo suggest that KAP1 sumoylation occurs on many sites simultaneously with the major ones located in the bromodomain. The PHD-Ubc9 interaction is required for this modification and potentially could also be regulated. Once modified by SUMO, KAP1 serves as a scaffold protein and recruits repression machinery through the recognition of the conjugated SUMO moieties by the SIM motifs of CHD3 and SETDB1, which recruit their associated proteins. Moreover, the KAP1-SETDB1 interaction not only results in recruitment of SETDB1 but also stimulates the HMTase activity of the enzyme. Ultimately, action of these effector molecules leads to chromatin reorganization at the promoter region, i.e. histone deacetylation and methylation of histone H3-K9. Notably, KAP1 scaffolding function is further exploited through recruitment of HP1. HP1 recognizes the H3-K9 methyl mark and participates in establishment of repressive chromatin state.

In summary, this study provides conclusive evidence that the KAP1 PHD finger acts as an intramolecular SUMO E3 ligase for the adjacent bromodomain and sumoylation of KAP1 as a major regulatory switch required for KRAB-KAP1 mediated repression.

REFERENCES

Boggio, R., Colombo, R., Hay, R. T., Draetta, G. F., and Chiocca, S. (2004). A mechanism for inhibiting the SUMO pathway. Mol Cell *16*, 549-561.

Capili, A. D., Schultz, D. C., Rauscher, I. F., and Borden, K. L. (2001). Solution structure of the PHD domain from the KAP-1 corepressor: structural determinants for PHD, RING and LIM zinc-binding domains. Embo J *20*, 165-177.

Chupreta, S., Holmstrom, S., Subramanian, L., and Iniguez-Lluhi, J. A. (2005). A small conserved surface in SUMO is the critical structural determinant of its transcriptional inhibitory properties. Mol Cell Biol *25*, 4272-4282.

Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X. P., Neilson, E. G., and Rauscher, F. J., 3rd (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev *10*, 2067-2078.

Gill, G. (2005). Something about SUMO inhibits transcription. Curr Opin Genet Dev *15*, 536-541.

Girdwood, D., Bumpass, D., Vaughan, O. A., Thain, A., Anderson, L. A., Snowden, A. W., Garcia-Wilson, E., Perkins, N. D., and Hay, R. T. (2003). P300 transcriptional repression is mediated by SUMO modification. Mol Cell *11*, 1043-1054.

Hecker, C. M., Rabiller, M., Haglund, K., Bayer, P., and Dikic, I. (2006). Specification of SUMO1- and SUMO2-interacting motifs. J Biol Chem *281*, 16117-16127.

Huntley, S., Baggott, D. M., Hamilton, A. T., Tran-Gyamfi, M., Yang, S., Kim, J., Gordon, L., Branscomb, E., and Stubbs, L. (2006). A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. Genome Res *16*, 669-677.

Joazeiro, C. A., and Weissman, A. M. (2000). RING finger proteins: mediators of ubiquitin ligase activity. Cell *102*, 549-552.

Johnson, E. S. (2004). Protein modification by SUMO. Annu Rev Biochem *73*, 355-382.

Kotaja, N., Karvonen, U., Janne, O. A., and Palvimo, J. J. (2002). PIAS proteins modulate transcription factors by functioning as SUMO-1 ligases. Mol Cell Biol *22*, 5222-5234.

Li, H., Ilin, S., Wang, W., Duncan, E. M., Wysocka, J., Allis, C. D., and Patel, D. J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature *442*, 91-95.

Lin, D. Y., Huang, Y. S., Jeng, J. C., Kuo, H. Y., Chang, C. C., Chao, T. T., Ho, C. C., Chen, Y. C., Lin, T. P., Fang, H. I., *et al.* (2006). Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors. Mol Cell *24*, 341-354.

Mukhopadhyay, D., and Dasso, M. (2007). Modification in reverse: the SUMO proteases. Trends Biochem Sci *32*, 286-295.

Peng, H., Begg, G. E., Schultz, D. C., Friedman, J. R., Jensen, D. E., Speicher, D. W., and Rauscher, F. J., 3rd (2000). Reconstitution of the KRAB-KAP-1 repressor complex: a model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein-protein interactions. J Mol Biol *295*, 1139-1162.

Ragvin, A., Valvatne, H., Erdal, S., Arskog, V., Tufteland, K. R., Breen, K., AM, O. Y., Eberharter, A., Gibson, T. J., Becker, P. B., and Aasland, R. (2004). Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300. J Mol Biol *337*, 773-788.
Rosendorff, A., Sakakibara, S., Lu, S., Kieff, E., Xuan, Y., DiBacco, A., Shi, Y., Shi, Y., and Gill, G. (2006). NXP-2 association with SUMO-2 depends on lysines required for transcriptional repression. Proc Natl Acad Sci U S A *103*, 5308-5313.
Ryan, R. F., Schultz, D. C., Ayyanathan, K., Singh, P. B., Friedman, J. R., Fredericks, W. J., and Rauscher, F. J., 3rd (1999). KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. Mol Cell Biol *19*, 4366-4378.
Sampson, D. A., Wang, M., and Matunis, M. J. (2001). The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. J Biol Chem *276*, 21664-21669.
Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G., and Rauscher, F. J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev *16*, 919-932.
Schultz, D. C., Friedman, J. R., and Rauscher, F. J., 3rd (2001). Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. Genes Dev *15*, 428-443.
Seet, B. T., Dikic, I., Zhou, M. M., and Pawson, T. (2006). Reading protein modifications with interaction domains. Nat Rev Mol Cell Biol *7*, 473-483.
Shi, X., Hong, T., Walter, K. L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Pena, P., Lan, F., Kaadige, M. R., *et al.* (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature *442*, 96-99.
Song, J., Zhang, Z., Hu, W., and Chen, Y. (2005). Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. J Biol Chem *280*, 40122-40129.
Sripathy, S. P., Stevens, J., and Schultz, D. C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. Mol Cell Biol *26*, 8623-8638.
Uchimura, Y., Ichimura, T., Uwada, J., Tachibana, T., Sugahara, S., Nakao, M., and Saitoh, H. (2006). Involvement of SUMO modification in MBD1- and MCAF1-mediated heterochromatin formation. J Biol Chem.
Wang, C., Ivanov, A., Chen, L., Fredericks, W. J., Seto, E., Rauscher, F. J., 3rd, and Chen, J. (2005). MDM2 interaction with nuclear corepressor KAP1 contributes to p53 inactivation. Embo J *24*, 3279-3290.
Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., Kauer, M., Tackett, A. J., Chait, B. T., Badenhorst, P., *et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86-90.
Yang, S. H., and Sharrocks, A. D. (2004). SUMO promotes HDAC-mediated transcriptional repression. Mol Cell *13*, 611-617.
Yang, X. J. (2004). Lysine acetylation and the bromodomain: a new partnership for signaling. Bioessays *26*, 1076-1087.

Zhou, Y., and Grummt, I. (2005). The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing. Curr Biol *15*, 1434-1438.

FIGURE LEGENDS

Figure 1. KAP1 is Sumoylated in Vivo and in Vitro in a PHD-Dependent Manner

(**A**) Sequence alignment of six KAP1 sumoylation sites. Forward and reverse sumoylation consensuses are shown and boxed.

(**B**) Schematic representation of KAP1 structural domains. B1 – box 1, B2 – box 2, CC – coiled-coil, HP1BD – HP1 binding domain. Amino acid positions of KAP1 sumoylation site lysines and various domain inactivating mutations are indicated. Deletion mutants of KAP1 utilized are abbreviated with the first letters of each domain (PB = PHD-Bromodomain). Sumoylation deficient mutant with all six sumoylation site lysines substituted to arginines is designated RR-R-RRR or K6R.

(**C**) Proteins were immunoprecipitated from H1299 cells with pre-immune serum or antibodies specific for different regions of KAP1: PB or RBCC; or an unrelated LimD1 antibody. A SUMO1 antibody was used in Western blot. Sumoylated forms of KAP1 are indicated with arrows.

(**D**) HEK293 cells were transfected with empty vector, Ubc9, SENP1 or dnSENP1 constructs with or without HA-SUMO1 plasmid as indicated. Proteins were detected by Western blot with KAP1 antibody. Sumoylated form of KAP1 is indicated with an arrowhead.

(**E**) In vitro sumoylation of GST-KAP1. Immobilized GST and GST-KAP1 were incubated in the presence or absence of recombinant E1, Ubc9, and ^{32}P-labeled SUMO1 as indicated. The reaction products were analyzed by coomassie blue staining (bottom) and autoradiography (top).

(**F**) HEK293 cells were transfected with the indicated FLAG-KAP1 and T7-SUMO1 plasmids. Proteins were analyzed by Western blot with FLAG antibody.

(**G-H**) In vitro sumoylation of indicated GST-KAP1 proteins. The major mono-sumoylated forms of KAP1 are indicated with arrowheads. Arrows indicate the position of major PHD-independent SUMO conjugates.

Figure 2. Mapping Sumoylation Sites and Repression Domains in KAP1

(**A**) In vitro sumoylation of GST fusions of indicated KAP1 domains wild type or various sumoylation site mutants was performed as in Fig.1E. The reaction products were analyzed by coomassie blue staining (bottom) and autoradiography (top). Arrowheads and arrows indicate the positions of mono- and di-sumoylated forms, respectively.

(**B**) HEK293 cells were transfected with plasmids expressing the same KAP1 domains as in (A) fused to LexA, together with 4xLexA-TK-luc plasmid and a pCMV–β-gal normalization vector.

(**C**) HEK293 cells were transfected with indicated LexA-PB plasmids, 4xLexA-TK-luc plasmid and a pCMV−β-gal normalization vector. The expression of LexA-PB proteins was confirmed by Western blot (bottom).

(**D**) HEK293 cells were transfected with T7-SUMO1 and indicated LexA-PB plasmids. LexA-PB proteins were immunoprecipitated with LexA antibody and analyzed by Western blot with SUMO1 (top) and LexA antibodies (bottom). Asterisk – IgGs.

(**E**) HEK293 cells were transfected with indicated LexA-PB plasmids and wild type or mutant Gam1 together with 4xLexA-TK-luc plasmid and a pCMV−β-gal normalization vector. The expression of LexA-PB and Myc-Gam1 proteins was confirmed by Western blot (bottom).

(**F**) In vitro sumoylation of indicated GST-NPB KAP1 proteins was performed as in Fig.1E. The reaction products were analyzed by coomassie blue staining (bottom) and autoradiography (top).

(**G**) HEK293 cells were transfected with HA-SUMO1 and indicated LexA-NPB plasmids. LexA-NPB proteins were immunoprecipitated with LexA antibody and analyzed by Western blot using HA (top) and KAP1 antibodies (bottom).

(**H**) HEK293 cells were transfected with the same LexA-NPB plasmids as in (G) together with 4xLexA-TK-luc plasmid and a pCMV−β-gal normalization vector. The expression of LexA-NPB proteins was confirmed by Western blot (bottom).

(B,C,E,H) Data are the mean $\pm$ SD of at least two experiments performed in duplicate.

Figure 3. The KAP1 PHD Binds to Ubc9 with High Selectivity and Directs Sumoylation of the Adjacent Bromodomain

(**A**) In vitro sumoylation of GST-PB wild type, indicated sumoylation site mutants and C651A mutant proteins. Arrowheads and arrows indicate the positions of mono- and di-sumoylated forms, respectively.

(**B**) HEK293 cells were transfected with indicated LexA-PB plasmids together with 4xLexA-TK-luc plasmid and a pCMV−β-gal normalization vector.

(**C**) GST and GST fusions of SUMO1, Ubc9, Sae1 and Sae2 were immobilized on glutathione beads and incubated with recombinant KAP1 PHD. Bound proteins were analyzed by coomassie blue staining. Lane 1 was loaded with a protein marker in addition to an input sample of the PHD protein.

(**D**) GST and GST-Ubc9 were immobilized on glutathione beads and incubated with the indicated recombinant KAP1 PHDs. Bound proteins were analyzed by coomassie blue staining.

(**E**) In vitro sumoylation of GST-PB modules from KAP1, TIF1α, TIF1γ and Sp100C was performed as in Fig.1E. The reaction products were analyzed by coomassie blue staining (bottom) and autoradiography (top).

(**F**) In vitro sumoylation of indicated GST fusion proteins: RING or RING with adjacent domain (RR) from mouse PIAS1 (aa 140-320), PHD or PB from KAP1, was performed as in Fig.1E. The reaction products were analyzed by coomassie blue staining (bottom) and autoradiography (top).

(**G**) In vitro sumoylation of GST-PHD-Bromo and GST-Bromo-PHD chimera was performed as in Fig.1E. The reaction products were analyzed by coomassie blue staining (bottom) and autoradiography (top).

Figure 4. Mapping of Ubc9 Binding Site on the KAP1 PHD Domain

(**A**) 2D ^{1}H-^{15}N HSQC spectral comparison of ^{15}N-labeled KAP1 PHD domain in the free form (black) and in the presence of Ubc9 (red). Molar ratio of KAP1 PHD finger to Ubc9 is 1:2.4. (Inset) Expansion of spectral region illustrating three peaks that undergo line broadening during Ubc9 titration. The blue peaks correspond to a mid-point of the titration, in which PHD:Ubc9 molar ratio is 1:1.2.

(**B**) Ribbon and surface representations of the KAP1 PHD domain structure (PDB code 1FP0), highlighting the residues that exhibited major resonance perturbations upon Ubc9 binding. KAP1 residues corresponding to NMR peaks that undergo line broadening at a PHD:Ubc9 molar ratio of 1:1.2 are in red, and those affected at 1:1.8 are in orange.

(**C**) Sequence alignment of PHD domains from KAP1 and its closest human PHD homologue BPTF as well as from TIF1 family proteins and Sp100C. Residues are color-coded according to ClastalW color scheme. The KAP1 PHD residues perturbed in NMR titration experiment (**A**,**B**)

are shaded in yellow, and aa substitutions which showed inhibitory effect in sumoylation assay (Figure S3) are shown underneath. Chelating cysteine and histidine residues are shaded in grey. Residues of two β-turn-β units of anti-parallel β-sheet are shaded in cyan.

(**D**) Sequence alignment of PHD domains from KAP1 and TIF1γ. The 6 aa substitutions of predicted loop segments L1, L2 and L3 are shown above. Identical residues are shaded in grey (chelating) and cyan.

(**E**) GST and GST-Ubc9 were immobilized on glutathione beads and incubated with the indicated recombinant KAP1 PHDs. Bound proteins were analyzed by coomassie blue staining.

(**F**) In vitro sumoylation of GST-PB wild type or indicated loop mutants was performed as in Fig.1E. The reaction products were analyzed by coomassie blue staining (bottom) and autoradiography (top).

(**G**) U2OS K4 cells were transfected and treated as in Fig.5C except that empty vector, FLAG-KAP1 wild type or indicated loop mutant plasmids were included into the transfection mixture.

Figure 5. Role of KAP1 Sumoylation in KRAB Domain Mediated Repression

(**A**) HEK293 cells were transfected with plasmids encoding LexA fusions of indicated repression domains together with empty vector, wild type or mutant Gam1, and 4xLexA-TK-luc plasmid

with a pCMV−β-gal normalization vector. The expression of the LexA fusion proteins (arrowheads) was confirmed by Western blot (bottom).

(**B**) HEK293 cells were transfected with FLAG-KAP1 and 100 ng of T7-SUMO1 plasmid together with empty vector, wild type or mutant Gam1 plasmids. Lysates were Western blotted with Myc antibody to detect Gam1 protein (bottom). Asterisk - non-specific band. KAP1 was immunoprecipitated with FLAG M2 antibody and Western blotted with FLAG (lanes 1-3), SUMO1 (lanes 4-6) and SUMO3 (lanes 7-9) antibodies.

(**C**) U2OS-G1 and -K4 cells were transfected with Gal4-DBD or Gal4-KRAB plasmids together with 5xGal4-TK-luc plasmid and a pTK−β-gal normalization vector. Endogenous KAP1 protein expression level was verified by Western blot (bottom).

(**D**) U2OS-K4 cells were transfected and treated as in (C) except that empty vector, FLAG-KAP1 wild type or indicated mutant plasmids were included into the transfection mixture (top). The expression of FLAG-KAP1 proteins was confirmed by Western blot (bottom).
(A,C,D) Data are the mean $\pm$ SD of at least two experiments performed in duplicate.

Figure 6. SETDB1 and CHD3 Interact with KAP1 in Yeast and in Vitro in SIM Dependent Manner

(**A,C,D**) Growth selection and quantitative β-gal analyses of L40a yeast cells transformed with indicated plasmids. The growth was scored on a three-point scale: (+++) - robust growth, (++) - slow growth and (-) – no growth on triple drop-out plates (right).

(**B**) L40a cells transformed with LexA-PB and -NPB constructs as in (A) were grown in double drop-out media. LexA fusion proteins were immunoprecipitated with LexA antibody and analyzed by Western blot with Smt3p (top) and LexA antibodies (bottom).

(**C**) Schematic representation of SETDB1 and CHD3 domain structure: T1 and T2 – Tudor domains, P – PHD, C – chromodomain. SIM is indicated by a stripped box.

(**E**) Sequence alignment of the SUMO interacting motifs (SIM) from SETDB1 and CHD3. The SIM consensus is boxed and the flanking acidic residues are italicized.

(**F**) Indicated GST fusion proteins immobilized on glutathione beads were incubated with ^{35}S-labeled SETDB1 or SETDB1dSIM. Bound proteins were analyzed by coomassie blue staining (bottom) and autoradiography (top). Brackets indicate positions of sumoylated protein forms.

(**G**) GST and GST fusions of SETDB1, SETDB1 SIMmt (aa 113-272), CHD3 (aa 1782-2000), CHD3dSIM (aa 1782-1994) and CHD4 (aa 1768-1912) immobilized on glutathione beads were incubated with recombinant unmodified or sumoylated full length KAP1. Bound proteins were analyzed by coomassie blue staining. Lanes 1 and 8 contain input samples.

Figure 7. KAP1 Sumoylation Promotes Recruitment of SETDB1 to the Target Promoter and Stimulates Its Enzymatic Activity

(**A**) U2OS-G1, -K4 and indicated reconstituted K4/FLAG-KAP1 cell lines were treated as in Fig.5C (top). The expression of KAP1, p53 and HP1α proteins was confirmed by Western blot (bottom).

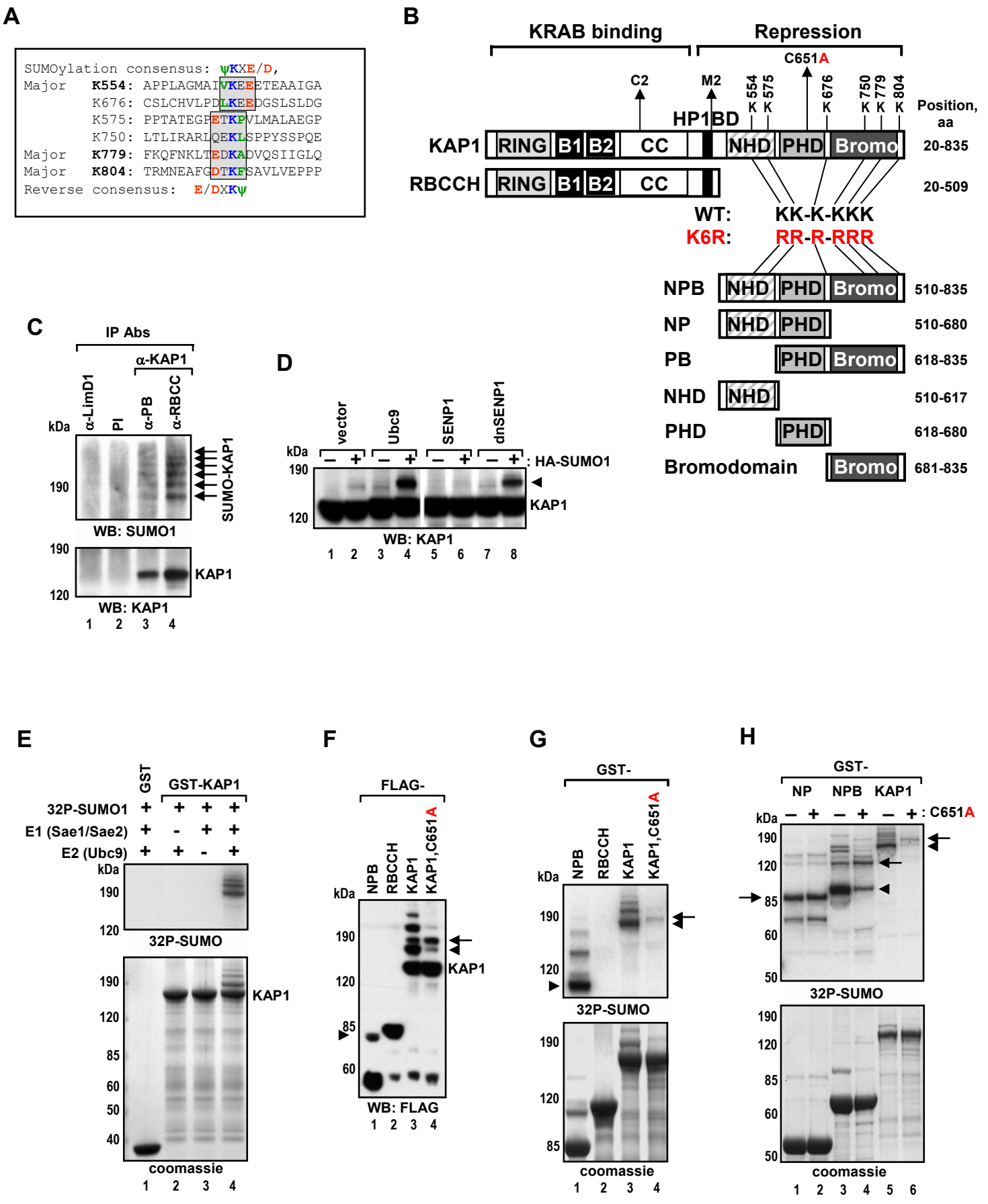
(**B**) Proteins from indicated reconstituted U2OS-K4/FLAG-KAP1 cell lines were immunoprecipitated with FLAG M2 resin in the presence or absence of 5 mM NEM and Western blotted with Mi2/CHD3, SETDB1, SUMO1, SUMO2 and FLAG antibodies.

(**C**) ChIP analysis of KAP1-mediated SETDB1 and HP1α recruitment to the target TK promoter. U2OS-K4 cells were transfected with Gal4-DBD or Gal4-KRAB, empty vector, FLAG-KAP1 wild type or indicated mutant plasmids together with 5xGal4-TK-luc plasmid.

(**D**) Increasing amounts of unmodified or sumoylated KAP1 PB were mixed with 100 ng of SETDB1 along with 5 μg of core histones and subjected to an in vitro HMTase assay. Autoradiograph shows corresponding ^{3}H-methyl-labeled histone H3. Western blotting confirms presence of SETDB1 and KAP1 PB proteins in the HMTase reaction. Arrowhead indicates the position of mono-sumoylated PB.

(**E**) KRAB-ZFP bound to the promoter of its target gene interacts with corepressor KAP1 via its N-terminal RBCC region. KAP1 PHD domain functions as an intramolecular E3 ligase by binding to Ubc9 (blue arrow) and directing sumoylation of the adjacent bromodomain. The SUMO-conjugated bromodomain recruits the CHD3/NuRD complex and SETDB1 through SUMO-SIM interaction, which results in the deacetylation of histones and the methylation of histone H3-K9. KAP1-bound HP1 recognizes H3-K9 methylation via its chromodomain (CD) and is deposited onto the promoter in order to silence gene expression. Small triangle – acetyl mark, small circle – H3-K9 tri-methyl mark on histone tails.

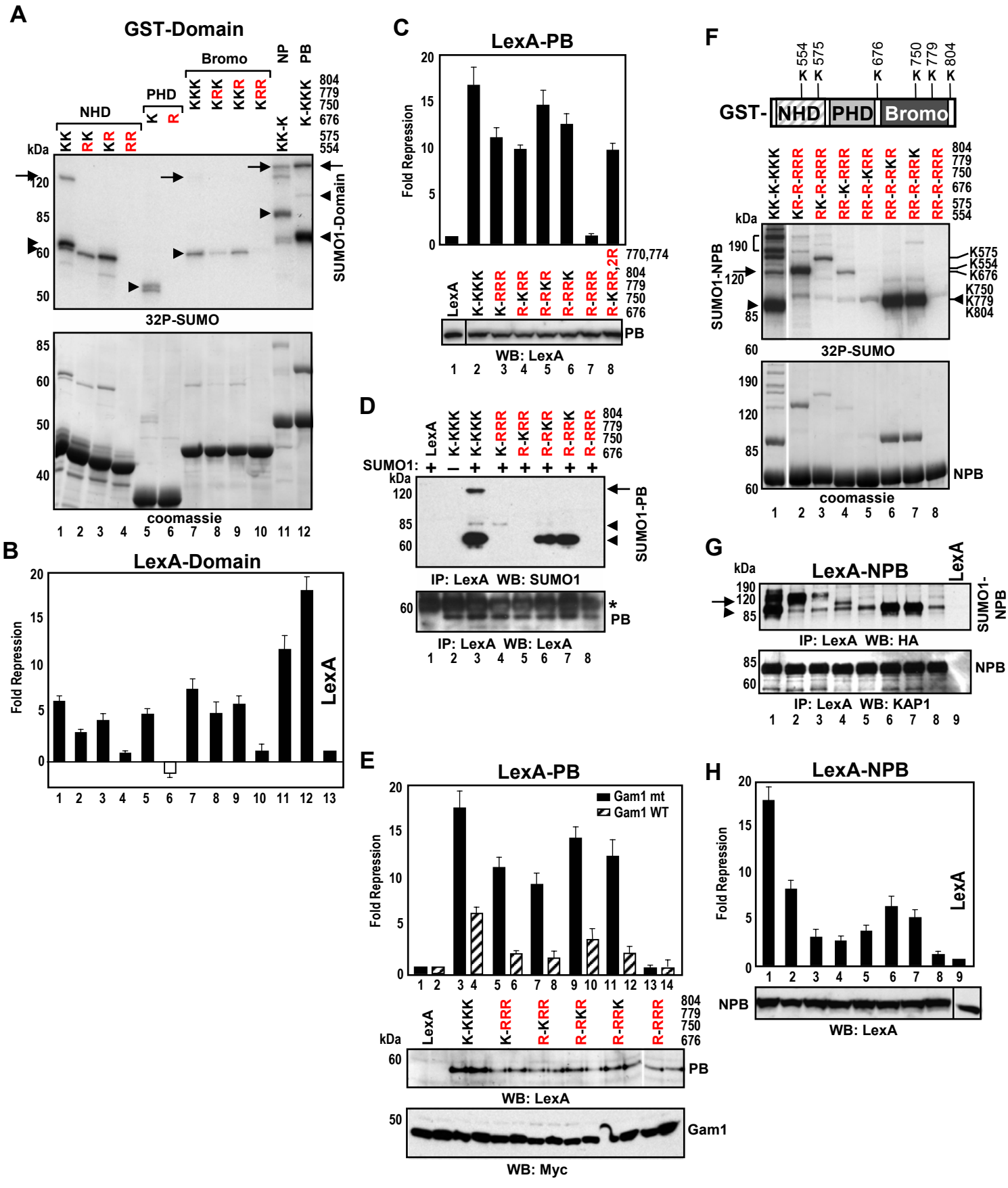
Fig.1 KAP1 is Sumoylated In Vivo and In Vitro in a PHD-Dependent Manner



Ivanov A. et al.

Figure 1

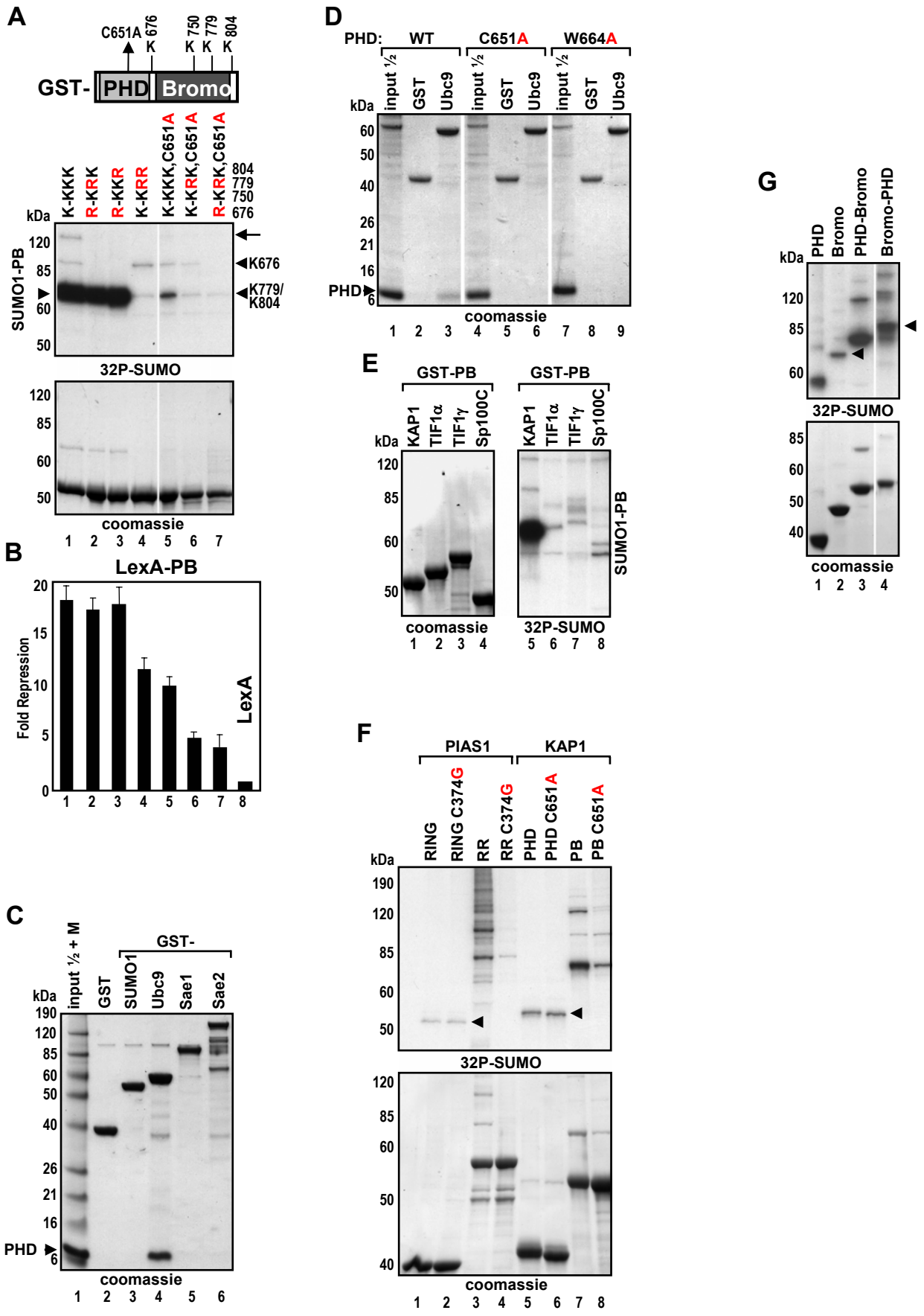
Fig.2 Mapping Sumoylation Sites and Repression Domains in KAP1



Ivanov A. et al.

Figure 2

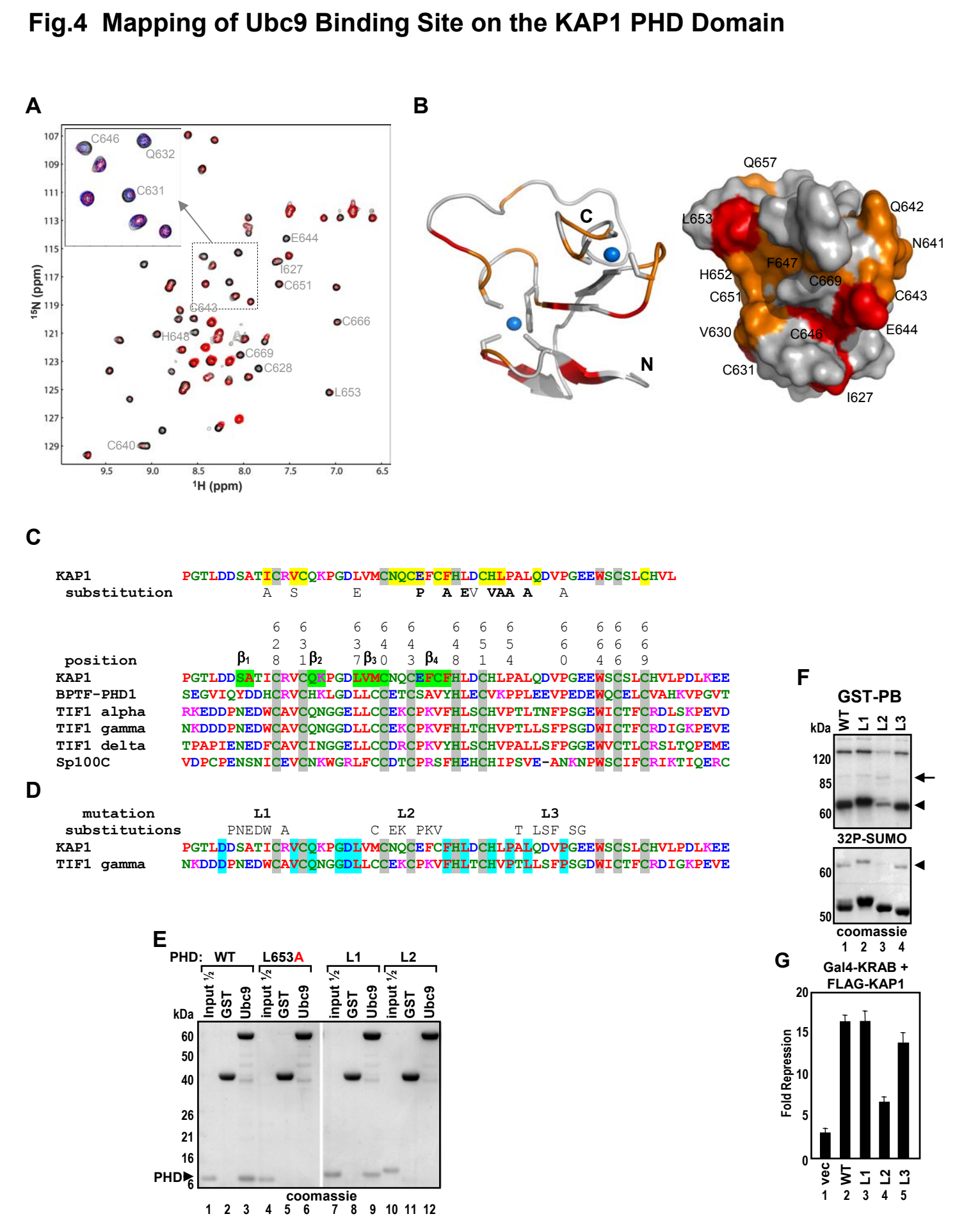
Fig.3 The KAP1 PHD Binds to Ubc9 with High Selectivity and Directs Sumoylation of the Adjacent Bromodomain



Ivanov A. et al.

Figure 3

Fig.4 Mapping of Ubc9 Binding Site on the KAP1 PHD Domain



Ivanov A. et al.

Figure 4

Fig.5 Role of KAP1 Sumoylation in KRAB Domain Mediated Repression

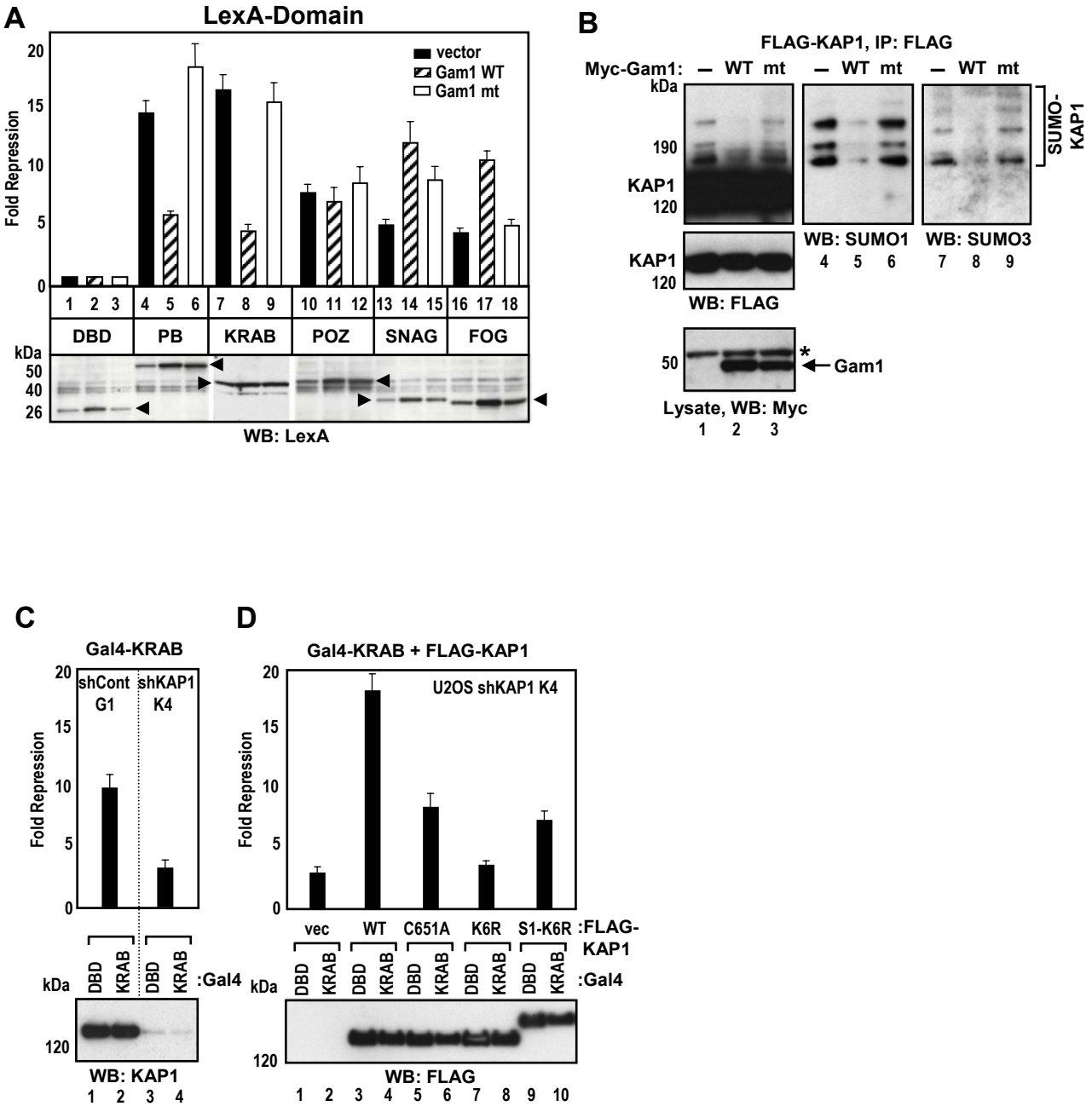
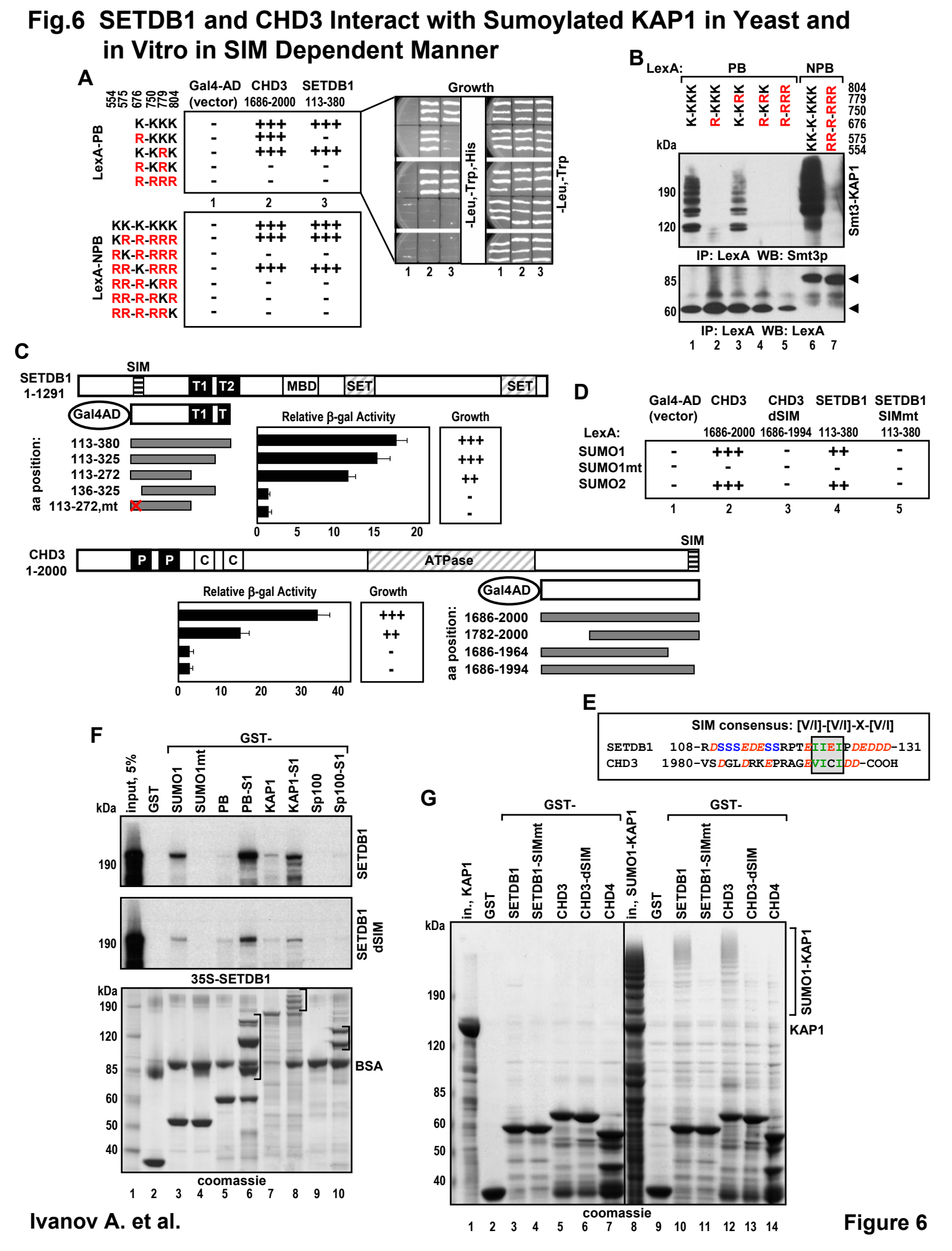


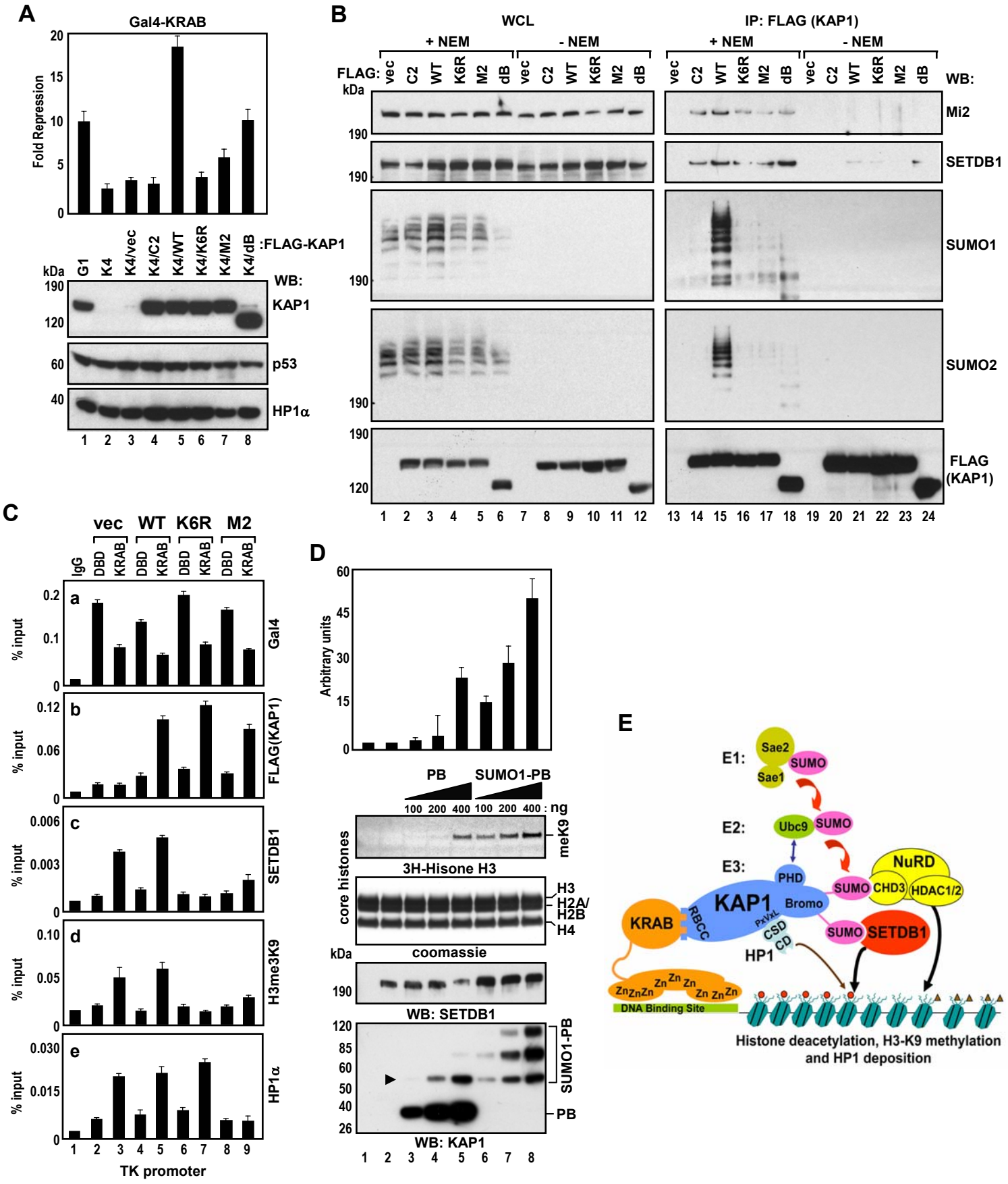
Fig.6 SETDB1 and CHD3 Interact with Sumoylated KAP1 in Yeast and in Vitro in SIM Dependent Manner



Ivanov A. et al.

Figure 6

Fig.7 KAP1 Sumoylation Promotes Recruitment of SETDB1 to the Target Promoter and Stimulates Its Enzymatic Activity



Ivanov A. et al.

Figure 7

The AJUBA LIM Domain Protein is a Co-Repressor for SNAG Domain Mediated Repression and Participates in Nucleo-Cytoplasmic Shuttling

Kasirajan Ayyanathan, Hongzhuang Peng, Zhaoyuan Hou, William J. Fredericks, Rakesh K. Goyal[§], Ellen M. Langer, Gregory D. Longmore[+] and Frank J. Rauscher, III*

The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104-4268

Key words: SNAG domain, LIM domain, Transcriptional Repression, Corepressors, Histone deacetylation, Nucleocytoplasmic signaling

*Corresponding author

Phone: 215-898-0995
FAX :215-898-3929
E-mail: rauscher@wistar.org

[+]Mailing address:

Department of Medicine
Washington University School of Medicine
660 South Euclid Avenue
St. Louis, MO 63110
Email: longmore@medicine.wustl.edu

[§]Mailing address:

Blood and Marrow Transplantation Program
Children's Hospital of Pittsburgh
3705 Fifth Avenue
Pittsburgh, PA 15213-2583
Email: goyalrk@pitt.edu

Abstract

The SNAG repression domain, named for its presence in the Snail/GFI-1 class of zinc finger transcription factors, is present in a variety of proto-oncogenic transcription factors and developmental regulators. The prototype SNAG domain containing oncogene, GFI-1 (growth factor independence-1) is responsible for development of T-cell thymomas. Significantly, these oncogenic functions require a functional SNAG repression domain. However, the molecular mechanism of SNAG domain mediated transcriptional repression that is responsible for these biological functions is completely unknown. Using a yeast two-hybrid strategy, we identified Ajuba and LIMD1 proteins, which can function as novel SNAG corepressors. Ajuba shuttles between the cytoplasm and the nucleus and may form a novel intracellular signaling system. Ajuba interacts with SNAG *in vivo*, co-localizes with it, and enhances SNAG-mediated transcriptional repression. We developed a strategy to recruit the mediators of SNAG repression to an expressed transgene in mammalian cells by using a synthetic, hormone-regulated SNAG repression domain. We show that in presence of hormone, the transgene is rapidly repressed due to the assembly of a multiprotein complex that comprises of Ajuba, LIMD1, possibly histone deacetylases and hitherto unidentified proteins at the promoter. We suggest that, in this molecular interaction, Ajuba may function as an adapter or molecular scaffold for the assembly of a macromolecular repression complex at the target promoters while the

SNAG domain may serve as a nuclear targeting motif in the transport of Ajuba into the nucleus.

Introduction

Generally, the transcription factors are modular in architecture; possess sequence-specific DNA binding-, and functionally separable effector-domains. The Cys2-His2 (C_2H_2) type zinc finger proteins (ZFPs) constitute a family of transcription factors, which can be classified into distinct sub-families such as SNAG-, KRAB-, and BTB/POZ-ZFP based on the highly conserved functional domains they possess at their amino termini. Mechanisms followed during transcriptional regulation include: 1) direct interaction with the components of basal transcription machinery, 2) ATP-dependent remodeling of the chromatin, and 3) site-specific modifications such as acetylation, phosphorylation, methylation and ubiquitination of the nucleosomal histones, which serve as biochemical codes in the context of complex chromatin structure. A widely accepted paradigm is that histone acetylation leads to gene activation while histone deacetylation results in repression. Often, the domains mediating the repression are proline-, alanine-, or glycine-rich (1). Importantly, many repressors function through corepressor molecules as exemplified by: DR-DRAP-1 (2), WRPW-Groucho (3), REST-Co-REST (4) and KRAB-KAP-1 (5).

SNAG repression domain (Snail/GFI-1) is present in the vertebrate homologues of the Snail/Slug and in the GFI-1 (growth factor independence-1) ZFPs (6-9). The *GFI-1* proto-oncogene was cloned by an insertional mutagenesis strategy wherein the Moloney murine leukemia virus-infected T-cells selected for IL-2-independence showed

non-random integration of the provirus at the *GFI-1* locus (10). Remarkably, *GFI-1* was independently cloned as a gene, which could cooperate with myc and pim1 in transgenic models of B- and T-cell lymphoma (11). The GFI-1 is a nuclear protein, recognizes a 12 bp consensus sequence (12). Potential target genes have been identified based on the presence of this recognition sequence in their promoter/enhancer regions (13). Full-length GFI-1 represses reporter plasmids containing these binding sites in multiple cell lines. Furthermore, the 20-amino acid SNAG domain alone fused to the zinc finger region was sufficient for repression and single amino-acid substitutions in the SNAG domain abolished this repression function (9). Consistent with its role as a T-cell tropic oncogene, overexpression of GFI-1 in immortalized T-cells allowed them to escape the G1 arrest induced by IL-2 withdrawal and, interestingly, a SNAG-domain mutant of GFI-1 was inactive in this capacity (9). Gene knockout studies have demonstrated that mammalian *GFI-1* genes are essential for the development of erythroid and megakaryocytic lineages (14, 15). A recent study has shown that GFI-1 is absolutely essential for neutrophil differentiation (16).

Several vertebrate homologues of the fly *Snail* and *Slug* genes have been cloned and their roles in development studied (17). In *Xenopus*, the Snail/Slug family has been shown to play essential roles in both mesoderm differentiation and in neural crest induction/migration and these functions require a competent SNAG domain. Both FGF and HGF induce the *Slug* gene's expression, which in turn induces epithelial-mesenchymal transitions (EMT) (18). The E2A-HLF fusion protein has been shown to activate the human *Slug* gene with resultant IL-3-independent survival and transformation of pro-B cells (19). Snail family members also play crucial roles in the development of

mesoderm and nervous system by triggering EMT. Recent studies have shown that the human and mouse E-cadherin promoters are direct targets for Snail ZFPs and the loss of E-cadherin expression is a major contributor to acquisition of an invasive, highly malignant phenotype during human tumor progression (20, 21). Thus, it is clear that the SNAG-ZFPs play important and diverse roles in embryonic development and in human disease. Despite these advances in defining the roles of SNAG domain proteins in biological processes, the molecular mechanisms of SNAG-mediated transcriptional repression are still completely unknown.

Using a yeast two-hybrid screen, we have identified Ajuba (22, 23) and LIMD1 (24) proteins as SNAG domain interactors. We have demonstrated that Ajuba functions as a SNAG domain-specific corepressor by performing thorough biochemical characterizations. We suggest that these LIM proteins mediate the repression function of SNAG-ZFPs by assembling a macromolecular repression complex at the target promoters.

Materials and Methods:

Cell lines and Antibodies: The COS-1, NIH3T3 and its stable cell line derivatives (SPHL11 and SPHL20) were grown as described (25). The α-PAX3 (26) and α-LEXA IgGs (Santa Cruz Biotechnology) are rabbit and goat polyclonal antibodies that recognize the PAX3 and LEXA DNA-binding domains (DBD), respectively. The α-Ajuba, α-LIMD1 are polyclonal antibodies made by immunizing rabbits with 6-His fusion proteins of mouse Ajuba (amino acids 1-216) and mouse LIMD1 (amino acids 1-158) as antigens. These antibodies do not cross-react with each other. The α-MYC tag (clone 9E10), α-

AcH3, α-AcH4 (raised against acetylated histone H3 & H4) and α-H3-MeK9 (raised against lys9-methylated histone H3) antibodies were purchased (Upstate).

Construction of a SNAG domain minigene and its derivatives: A synthetic SNAG domain coding sequence derived from *GFI-1* was constructed by overlap-extension-PCR. A Kozak consensus was added to the initiator methionine of the 20-amino acid SNAG domain (**MPRSFLVKSKKAHSYHQPRS**) plus a six amino acid spacer (**PGPDYS**) and in-frame EcoR I and Sal I sites were included for cloning into pSP73 vector. Next, we performed alanine-scanning mutagenesis for the analysis of the SNAG repression domain. Using appropriate mutant oligos, we first changed the FLV to AAA and KK to AA; later mutations involved single amino acid substitutions S-A and F-A as illustrated in **Figures 1C** and **2A**. These wild type (wt) and mutant SNAG minigenes were used to construct PAX3-, LEXA-, and pBTM116-LEXA fusions.

Immunoprecipitation (IP) and Electrophoretic Mobility Shift Assays (EMSA): Each expression plasmid was verified for stable protein expression *in vivo* as described (25). Briefly, [^{35}S]-L-methionine labeled whole cell extracts prepared from transiently transfected COS-1 cells were immunoprecipitated with α-PAX3 IgG or α-LEXA IgG, and analyzed by SDS-PAGE and fluorography. In co-immunoprecipitation experiments, subsequent to co-transfection of either SNAG-PAX3 or SNAG-LEXA with MYC-Ajuba plasmid (1:1 ratio), immunoprecipitation was carried out in mild conditions using ELB buffer (27). Nuclear extracts were prepared using a rapid protocol (26) and used in EMSA with ^{32}P-labeled e5 (PAX3 site derived from the *engrailed* gene) probe.

Transcriptional repression assays: To monitor the repression potentials of the chimeric SNAG repressor proteins, 2 x 10^5 NIH3T3 cells were transfected with 1 μg of the

expression plasmid along with 0.5 μg of PAX3- or LEXA-luciferase reporter plasmids, and 0.25 μg of pCMV-LacZ plasmids, using lipofectAMINE (Life Technologies). Whole cell extracts were assayed for luciferase activities and then normalized to β-galactosidase values for transfection efficiency (25, 26). Fold repression was determined as the ratio of normalized light units in vector versus SNAG-expression plasmid transfected cells. To examine the effect of Ajuba on SNAG-PAX3 or SNAG-LEXA-mediated repression, co-transfection was performed as above.

Immunofluorescence and co-localization: Immunocytochemistry was performed essentially as described (27). NIH3T3 cells were transfected with the indicated expression plasmids, fixed, permeabilized, and then incubated with α-LEXA (1:100 dilution) or α-MYC (1:1000 dilution) antibodies. The immunostained cells were detected with FITC-conjugated α-rabbit IgG (1:500 dilution) or Texas Red-conjugated α-mouse IgG (1:1000 dilution), respectively. Finally, the cells were stained for DNA with Hoechst (2 ng/ml) and mounted on glass slides using Fluoromount G (Southern Biotechnology Associates, Inc). Cells were visualized using a Leica confocal laser-scanning microscope. Two channels were recorded simultaneously if no cross talk could be detected. Images were captured using QED Imaging software.

Chromatin Immunoprecipitation (ChIP): SPHL11 or SPHL20 cells were plated at 5 x 10^6 cells/150-mm dish, treated continuously with either 500 nM 4-OHT (+OHT dishes) or 0.1% ethanol (-OHT dishes) for 2 days, and then fixed in 1% formaldehyde (EM Biosciences) for 20 min at 37^0C. Solubilized, sonicated chromatin was prepared as described (28), and immunoprecipitated with α-PAX3, α-Ajuba, α-LIMD1, α-AcH3, α-AcH4 or α-H3-MeK9 antibodies. Usually, 10% of the clarified chromatin was saved as

input. The immune complexes were processed as described (28). Both the input and the immunoprecipitated DNAs were used in quantitative PCR reactions with the primer-pairs illustrated in **Figure 5A**. The **PBS1** (5' GATCGATAATTCGAGCTACTG 3') and **PBS2** (5' GAGCTCGGTACCCGGGTCG 3') primer-pair amplify the PAX3 binding sites; the primer-pair **TKP1** (5' GCGCGGTCCCAGGTCCACTT 3') and **LUC1 (**5' TCCAGGAACCAGGGCGTATCTCT 3') amplify the HSV-TK promoter region. The DNA fragments were electrophoresed on 1.5% agarose gels and either photographed or Southern-blotted and autoradiographed.

Results

Characterization of engineered SNAG repressors: SNAG transcription factor superfamily includes Snail, Slug, GFI-1, Scratch, and IA-1 ZFPs. In spite of possessing a varied number of C_2H_2-type zinc fingers at their COOH-termini, they all contain a highly conserved NH_2-terminal SNAG repression domain (**Figure 1A**). Multiple sequence alignment using Clustal W indicates that the SNAG domain homology extends to ~20 amino acids with the NH_2-terminal first seven amino acids being strictly conserved between the different members (**Figure 1B**). Further, the SNAG domains of *GFI-1* and *GFI-1B* genes also possess a well-conserved nuclear localization signal (KSKK). To determine if the synthetic SNAG domain would function as a transferable, modular repression domain, we fused it to the NH_2 terminus of the minimal PAX3 DNA binding domain (PAX3) (**Figure 1C**). We have used the PAX3 as a recipient for the following reasons, the PAX3: 1) binds DNA as a monomer and recognizes an extended non-degenerate DNA binding site due to paired- and homeodomain-DNA binding motifs, 2)

is easily detectable using PAX3 antibodies, and 3) is essentially neutral when bound to DNA in the absence of an effector domain. We used the alanine-scanning mutagenesis to map the amino acid sequence requirement for active repression. We targeted both the extreme NH_2 terminal block of seven amino acids, which is identical among all SNAG domains and the more COOH-terminal region, which is not as highly conserved. After PCR-mediated site-directed mutagenesis and DNA sequence analysis, they were sub-cloned into the PAX3 constructs (**Figure 1C**). Each protein was comparably expressed (**Figure 1D**) and showed DNA binding activity to e5 site to the same extent (**Figure 1E**). Expression plasmids encoding SNAG-PAX3 fusions were co-transfected with the PAX3-luciferase reporter plasmid and cell extracts assayed for activity. As depicted in **Figure 1F**, fusion of the SNAG domain to the NH_2-terminus of PAX3 created a potent repressor. Repression was dose-dependent and was strictly contingent upon PAX3 binding sites in the reporter plasmid (data not shown). However, the mutants varied in their abilities to repress the PAX3-luciferase in reporter assays (**Figure 1F**). The FLV-AAA substitution completely abolished the repression. However, the S-A and KK-AA mutants retained significant repression activity whereas the single substitution F-A abolished repression function. Thus, phenylalanine to alanine substitution in the highly conserved SNAG domain converts a normally powerful repressor of transcription into a neutral DNA binding protein. Whether the seven-amino acid NH_2 terminal segment is sufficient for repression remains to be tested. Interestingly, a modular repression domain as small as 4 amino acids (WRPW) present in the hairy and achaete-scute bHLH transcription factors is sufficient to confer repression by recruiting the Groucho corepressor (3). We suggest that the SNAG domain may function in a similar manner.

A Yeast Two-Hybrid Strategy to identify SNAG domain-Associated Proteins (SNAPS): We reasoned that, like other repression domains, direct protein-protein interactions might be required for SNAG mediated repression. One of the strategies for isolating a SNAG-specific corepressor is to perform a yeast-two hybrid screen using the 20-amino acid SNAG domain as bait and to counter-screen the preys with the SNAG domain mutants that lack repression function. We generated a SNAG-LEXA fusion by placing the SNAG domain at the NH_2-terminus of the LEXA DNA binding domain (**Figure 2A**). In order to limit the spectrum of potential protein-protein interactions critical for SNAG function, we only used the 20-amino acid SNAG domain and not the full NH_2 terminus of GFI-1. However, in order to ensure that there were no spatial constraints to protein binding we included a small spacer of 17 amino acids of PAX3 between the SNAG domain and the beginning of the LEXA. We also constructed mutant SNAG-LEXA fusions with two mutants that abolished repression, and two mutants, which showed wild-type levels of repression activity (**Figure 2A**). The SNAG-LEXA fusions were properly expressed in mammalian cells (**Figure 2A**), exhibited proper binding to a LEXA DNA-binding site in gel shift assays (data not shown), and repressed the LEXA-luciferase reporter plasmids (**Figure 2A**). Thus, this spectrum of "baits" formed the important set of reagents with which to identify and distinguish true mediators of SNAG repression.

First, we showed that the pBTM116-SNAG-LEXA constructs neither activated nor repressed transcription in yeast (data not shown). The wt SNAG-LEXA fusion was then co-transformed with a murine embryonic (E9-E11) library constructed in pVP16 and scored for histidine prototrophy and LacZ activity in a standard two-hybrid screen. From

~ 200 primary hits that were obtained upon screening ~40 million library clones, we isolated ~20 that interacted strongly with the wt SNAG domain. Several of these were re-screened by reintroduction into naive yeast with a variety of other baits. As shown in **Figure 2B**, SNAP13 and SNAP20 interacted very strongly with wt SNAG but failed to interact with the SNAG (FLV-AAA) mutant, LEXA, or any of the other negative control baits including KAP-1, Lamin, Rho and PKC. We further characterized SNAP13 and SNAP20 by reintroducing them into naive yeast along with each of the pBTM116-SNAG-LEXA baits (mating assay). As shown in **Figure 2C**, their interactions were almost completely concordant with repression activity; each interacted with active SNAG repressors (wt and S-A) but did not interact with the mutants (FLV-AAA and F-A) which abolish repression. The exception to this concordance is KK-AA, which is expressed at a reduced level in yeast (data not shown). Thus, by these criteria these clones are strong candidates for mediators of SNAG function.

We rescued the inserts, and the deduced DNA sequence was in-frame with the VP16 activation domain encoded by the vector. In the primary screen, we obtained two independent hits of SNAP13 and multiple but identical isolates of SNAP20. BLAST searches revealed that SNAP13 and SNAP20 were identical to the two NH_2-terminal LIM domains of Ajuba and LIMD1, respectively. LIM is an acronym of three transcription factors, Lin11, Isl-1, and Mec-3, in which the motif was first identified (29). The LIM domain protein sub-family that includes Ajuba and LIMD1 is illustrated in **Figure 2D**. The Ajuba, LIMD1, and other three-LIM domain proteins such as Zyxin, TRIP6, LPP contain characteristic glycine/proline-rich regions and nuclear export signals (NES). They belong to Group 3 LIM domain family whose members exhibit nucleocytoplasmic

distribution and mediate both cellular and nuclear signaling events. LIM domains are bona fide protein-protein interaction motifs, which encode a signature Cys2-His1-Cys3-Cys Zn^{2+} binding domain as shown for the three individual LIM domains of Ajuba (**Figure 2E**). A dendrogram analysis constructed from the LIM domains of the three LIM domain proteins revealed that a distinct phylogenetic relationship exists only between the LIM domains of Ajuba and LIMD1 (**Figure 2F**). The SNAG-interacting LIM domains of Ajuba and LIMD1 (about 134 amino acids) fell into a single group with a similarity index of 62.7, implying that they evolved from a common ancestral gene.

Ajuba interacts with SNAG *in vivo* and functions as a corepressor: To verify whether the SNAG-LIM domain interactions occur *in vivo* in mammalian cells, we used expression vectors for SNAG-PAX3, SNAG-LEXA, and MYC-epitope tagged full length Ajuba (MYC-Ajuba). Proper expression of the MYC-Ajuba protein was confirmed by immunoprecipitation with α-MYC MAb (data not shown). Next, we co-transfected the MYC-Ajuba and wt or mutant SNAG-LEXA plasmids into COS-1 cells. The ^{35}S-methionine-labeled cell lysates were immunoprecipitated with either α-MYC or α-LEXA IgG. **Figure 3A** shows that the α-MYC detects a 66-kDa MYC-Ajuba protein while the α-LEXA detects a 25-kDa LEXA protein in the MYC-Ajuba+LEXA vector co-transfected cells. However, in the wt SNAG-LEXA+MYC-Ajuba co-transfected cells, the α-LEXA also detects the MYC-Ajuba protein. As expected, the α-MYC also detects the wt SNAG-LEXA protein in the same cell population. Strikingly, similar co-immunoprecipitation experiments with mutant SNAG-LEXA plasmids indicated that the repression-incompetent mutants of the SNAG domain fail to bind the Ajuba, strongly suggesting that this interaction is functionally relevant (**Figure 3A**).

To confirm that the region of Ajuba required for SNAG interaction contained the LIM domain, we used the expression plasmids coding for the NH_2- or COOH-terminal segments of Ajuba (defined as "Ajuba-PreLIM" or "Ajuba-LIM") in co-transfection experiments along with the SNAG-PAX3 plasmid. We observed that the α-PAX3 sera detects the SNAG-PAX3 and also co-precipitates the Ajuba protein or the Ajuba-LIM domain but not the Ajuba-preLIM domain, proving that the LIM region of Ajuba is the site of SNAG domain interaction (data not shown). Thus, Ajuba-SNAG-PAX3 complexes are very efficiently formed *in vivo* following co-transfection and are stable to moderate salt, detergent extraction procedures. Importantly, this co-precipitation was not detected when a PAX3 gene that lacks a SNAG domain was used, confirming that a SNAG-Ajuba interaction is occurring (data not shown).

We performed co-transfection experiments to determine if Ajuba can modify the transcriptional repression function of the SNAG domain. We transfected CMV vectors for SNAG-PAX3 or SNAG-LEXA constructs into NIH3T3 cells, with or without CMV-Ajuba. SNAG-LEXA repressed the LEXA-luciferase reporter, and this repression was enhanced by CMV-Ajuba. As expected, Ajuba also enhanced SNAG-PAX3 mediated repression of the PAX3-luciferase reporter (**Figure 3B**). These promising results suggest that Ajuba mediates SNAG repression and functions as a SNAG corepressor.

Recruitment and Nuclear Co-localization of Ajuba and SNAG in NIH3T3 cells: The functional studies prompted us to investigate the sub-cellular localization patterns of the MYC-Ajuba, the wt and mutant SNAG-LEXA proteins. Immunofluorescence analysis revealed that cells transfected with MYC-Ajuba showed predominant cytoplasmic staining and those that were transfected with SNAG-LEXA (wt and mutants) showed

intense nuclear staining (**Figure 4A**). However, cells that received both Ajuba and SNAG-LEXA plasmids showed abundant nuclear staining of Ajuba that was solely dependent on the repression competency (**Figure 4B**). A majority of cells in this doubly transfected population showed the pattern observed in **Figure 4B.** These remarkable images strongly corroborate our functional studies described earlier. These results also suggest that the SNAG domain can recruit and/or retain Ajuba in the nucleus, and are consistent with our earlier observations that Ajuba is both cytoplasmic and nuclear. Recent reports suggest that Group 3 LIM proteins like Zyxin and Ajuba can shuttle between the cytoplasm and the nucleus and thus may form a novel intracellular signaling system. It has been hypothesized that there are nuclear "anchors" for this class of shuttling LIM proteins (30). We hypothesize and present evidence that the SNAG domain may be such an anchor in case of Ajuba and LIMD1.

Molecular components of SNAG repression are recruited to the target promoter: We investigated whether the SNAG domain mediated transcriptional repression involves chromatin modification by using trichostatin-A (TSA), a specific inhibitor of histone deacetylase (HDAC) enzyme. We generated stable cell lines, which contain a chromatin-integrated PAX3-luciferase reporter transgene and a conditional SNAG repressor (SNAG-PAX3-HBD) (SPHBD) (**Figure 5A**). We tested multiple cell lines for protein expression and 4-OHT dependent repression of the PAX3-luciferase reporter (25) (data not shown). We chose the SPHL11 and SPHL20 clones to study the effect of TSA due to their repression potentials. In presence of 4-OHT, about six-fold repression was observed, however, in presence of both 4-OHT and TSA, this repression was totally relieved in both clones (data not shown). This experiment suggests that one of the

mechanisms of SNAG domain repression involves chromatin modification. To confirm these biochemical evidences of SNAG repression mechanisms at the molecular level, we performed comprehensive chromatin immunoprecipitation (ChIP) experiments using a battery of antibodies raised against putative molecular components. Chromatin-associated proteins were cross-linked to DNA *in vivo* with formaldehyde in mock or 4-OHT treated SPHL11 and SPHL20 cells and immunoprecipitated with α-PAX3, α-Ajuba, α-LIMD1, α-AcH3, α-AcH4, and α-MeK9 antibodies. The immunoprecipitated DNA was analyzed by quantitative PCR, using primer pairs for the 6xPAX3 binding sites (PBS1&PBS2), or the TK promoter regions (TKP1&LUC1) of the CD19-TK-LUC Zeo^R locus. Fragments, which bracket the PAX3 binding sites, were considerably higher in the PAX3 immunoprecipitates (IPs) after 4-OHT-treatment indicating that there was a strong enrichment of the SPHBD protein at the PAX3 sites. This experiment was controlled by the presence of similar levels of PAX3 sites in the input chromatin (**Figure 5B**). Likewise, we observed a strong, 4-OHT-dependent enrichment of the 257 bp TK promoter fragment in a-Ajuba, a-LIMD1, and a-MeK9 chromatin IPs (**Figure 5C**). Thus, in addition to the DNA binding component, other components of the SNAG repression complex (i.e. Ajuba, LIMD1 and H3-MeK9) were inducibly recruited to the target gene. Based on these results, we conclude that the SPHBD fusion protein, and Ajuba, LIMD1, H3-MeK9 are strongly recruited to the chromatin regions comprising the PAX3 binding site, and the TK promoter region, respectively.

In parallel ChIP experiments, we observed significant reduction in the levels of acetylated histone H3 and H4 at the TK promoter region upon 4-OHT-treatment, indicating that the histone hypoacetylation occurs upon recruitment of SPHBD protein to

the upstream PAX3 sites (**Figure 5D**). These results are in agreement with the complete reversal of SNAG repression we observed in presence of TSA. Future chromatin immunoprecipitation experiments will unravel the identity of the HDAC involved in SNAG repression. We expect these experiments to yield clues about the nature of the unique nucleocytoplasmic signaling mechanism operated by Ajuba and LIMD1 proteins.

Discussion

Mammalian zinc finger transcription factors play vital roles during organism development and subsequent homeostasis. In order to gain insights into the molecular mechanisms of SNAG-mediated repression that are responsible for the biological functions of SNAG-ZFPs, we isolated and characterized SNAG-specific corepressors. The modular nature of SNAG and other repression domains such as KRAB, BTB/POZ aids to isolate them away from the other functional elements present in the full-length proteins and study them in detail. We followed this approach in our earlier studies to understand the mechanistic details of the KRAB mediated repression by isolating and characterizing the KAP-1 (KRAB Associated Protein-1) corepressor (5, 27). In this study, we have identified Ajuba and LIMD1 proteins as downstream players in the SNAG repression pathway.

Subsequent to the initial identification of LIM domain proteins Ajuba and LIMD1 as SNAG-domain interactors in the yeast two-hybrid assay, we performed extensive biochemical analyses to confirm this interaction. Our striking results from co-immunoprecipitation experiments indicate that SNAG domain interacts avidly with Ajuba *in vivo* and also that Ajuba can enhance SNAG-mediated repression. Interestingly,

we observed these results only with repression-competent SNAG derivatives. The KAP1 corepressor, which interacts only with repression-competent KRAB domains also enhance KRAB-mediated repression. Since, Ajuba possesses these important characteristics of corepressor proteins similar to KAP1, we designate Ajuba as a SNAG-corepressor.

LIM domain interactions have been observed in several classes of transcription factors and cofactors. Specific examples include: 1) The nuclear LIM-only (LMO) protein interact with GATA-1, 2 and Tal1/Scl transcription factors (31); 2) The LIM-domain binding protein Ldb1, its LIM-only protein partner LMO2, and the DNA-binding SCL/E12 exist in a multiprotein complex that negatively regulates erythroid function (32); 3) FHL2/DRAL interacts with the promyelocytic leukemia zinc-finger (PLZF) protein (a sequence-specific transcriptional repressor) and functions as a corepressor by augmenting PLZF-mediated transcriptional repression (33); 4) FHL2 associates with Jun and Fos and serves as a co-activator of AP-1 by stimulating Fos- and Jun-dependent transcription (34); 5) The CLIM forms a protein complex consisting of LIM-HD, LMO, bHLH, GATA, at the target promoters (35); 6) The LIM-homeodomain transcription factor Lhx3 binds to the pituitary-specific transcription factor Pitx-1 and also interacts with several co-activator/adapter proteins including the SLB, selective LIM-binding protein (36); and 7) Cysteine-rich LIM-only proteins (CRP1 and CRP2) play important roles in organizing multiprotein complexes, both in cytoplasm, where they participate in cytoskeletal remodeling, and in the nucleus, where they strongly facilitate smooth muscle differentiation (37). These studies clearly demonstrate that nuclear LIM proteins can

function as adapter molecules in the formation of large multiprotein complexes that form on DNA and that influence transcription.

Our co-localization studies demonstrate that Ajuba by itself remains totally cytoplasmic, however, in the presence of repression competent SNAG domains, its localization completely changed to nuclear. We suggest that SNAG domain facilitates both import of Ajuba into the nucleus and its subsequent retention there by serving as a nuclear anchor. Similar nuclear targeting functions have been observed for other LIM proteins. For example: 1) Stimulation of the Rho signaling pathway induces translocation of the transcriptional LIM-only coactivator FHL2 to the nucleus (38); 2) In heart, FHL2 interacts with hNP220, a DNA-binding nuclear protein to serve as a molecular adapter in the formation of a multiprotein complex (39); 3) The LIM protein KyoT2, an alternatively spliced murine isoform of SLIM1 has been shown to negatively regulate transcription by interaction with RBP-J DNA binding protein and displacing it from the DNA (40, 41); 4) Interaction of Zyxin (usually found in focal adhesions) with the human papillomavirus-derived E6 protein leads to its accumulation in the nucleus where it functions as a transcriptional activator. Significantly, this interaction requires the three LIM domains present at the carboxyl terminus of Zyxin (42); and 5) The Lipoma-Preferred Partner (LPP) protein exhibits nucleocytoplasmic distribution and possesses focal adhesion and nuclear targeting capabilities. It does not contain any consensus nuclear localization signals suggesting that it may be imported into the nucleus via a nuclear localization signal containing transport protein (43).

The co-localization studies also suggested that upon SNAG-mediated translocation, Ajuba along with its (unidentified) associated proteins could constitute a

macromolecular repression complex at the target promoter. Based on our findings from the chromatin immunoprecipitation experiments, we favor the notion that the initial DNA-binding by the SNAG-PAX3 repressor at the PAX3 binding sites leads to the subsequent recruitment of Ajuba, and their associated proteins to the promoter region. Since we also observed significant hypoacetylation of histone tails at the promoter region and complete reversal of SNAG-mediated repression in presence of TSA, we strongly believe that SNAG repression involves histone deacetylases and that HDAC may be a constituent in the SNAG-holo repression complex.

At present, there are nine HDAC family members belonging to three distinct families. The class I HDAC family members, which consists of HDAC 1, 2, 3 and 8, are nuclear while the class II members, which consists of HDAC 4, 5, 6 and 7, are actively maintained in the cytoplasm and are imported to the nucleus when required. The class III consists of NAD^+-dependent histone deacetylases. In contrast to the class II, and I members of the class III histone deacetylases are insensitive to the histone deacetylase inhibitor trichostatin-A. Based on our observations, we hypothesize that either Ajuba interacts with a cytosolic HDAC and recruits it to the nucleus, or it interacts with a HDAC existing in the nucleus and recruits it to the repression complex. In support of this, RLIM has been shown to interact with members of histone deacetylase corepressor complex (44). It would be interesting to identify the nature of the HDAC involved, and to decipher the hitherto unidentified associated proteins that constitutes the macromolecular repressor complex. This necessitates isolation and characterization of interacting proteins that are present in both cytoplasmic and nuclear compartments.

Acknowledgements

We thank S. Hollenberg for the murine embryonic (E9-E11) library. We thank D.C. Schultz for the LEXA-luciferase reporter plasmid. K.A. is supported by NIH (KO1-CA095620). G.D.L. is supported by the American Heart Association (9940116N), the Edward Mallinckrodt Jr. Foundation, and the NIH (RO1-CA75315). G.D.L. is an Established Investigator of the American Heart Association. F.J.R. is supported in part by National Institutes of Health grants, Core grant CA10815, CA92088, CA095561, DAMD17-96-1-6141, 17-02-1-0631, the Irving A. Hansen Memorial Foundation, the Susan G. Komen Breast Cancer Foundation, and the Emerald Foundation. Support from Commonwealth Universal Research Enhancement Program, Pennsylvania Department of Health is gratefully acknowledged.

References:

1. Saha, S., Brickman, J. M., Lehming, N. & Ptashne, M. (1993) *Nature* 363, 648-52.
2. Kim, S., Na, J. G., Hampsey, M. & Reinberg, D. (1997) *Proc Natl Acad Sci U S A* 94, 820-5.
3. Fisher, A. L., Ohsako, S. & Caudy, M. (1996) *Mol Cell Biol* 16, 2670-7.
4. You, A., Tong, J. K., Grozinger, C. M. & Schreiber, S. L. (2001) *Proc Natl Acad Sci U S A* 98, 1454-8.
5. Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X. P., Neilson, E. G. & Rauscher, F. J., 3rd (1996) *Genes Dev* 10, 2067-78.

6. Hemavathy, K., Guru, S. C., Harris, J., Chen, J. D. & Ip, Y. T. (2000) *Mol Cell Biol* 20, 5087-95.

7. Nakakura, E. K., Watkins, D. N., Schuebel, K. E., Sriuranpong, V., Borges, M. W., Nelkin, B. D. & Ball, D. W. (2001) *Proc Natl Acad Sci U S A* 98, 4010-5.

8. Nibu, Y., Zhang, H., Bajor, E., Barolo, S., Small, S. & Levine, M. (1998) *Embo J* 17, 7009-20.

9. Grimes, H. L., Chan, T. O., Zweidler-McKay, P. A., Tong, B. & Tsichlis, P. N. (1996) *Mol Cell Biol* 16, 6263-72.

10. Gilks, C. B., Bear, S. E., Grimes, H. L. & Tsichlis, P. N. (1993) *Mol Cell Biol* 13, 1759-68.

11. Zornig, M., Schmidt, T., Karsunky, H., Grzeschiczek, A. & Moroy, T. (1996) *Oncogene* 12, 1789-801.

12. Zweidler-Mckay, P. A., Grimes, H. L., Flubacher, M. M. & Tsichlis, P. N. (1996) *Mol Cell Biol* 16, 4024-34.

13. Grimes, H. L., Gilks, C. B., Chan, T. O., Porter, S. & Tsichlis, P. N. (1996) *Proc Natl Acad Sci U S A* 93, 14569-73.

14. Karsunky, H., Zeng, H., Schmidt, T., Zevnik, B., Kluge, R., Schmid, K. W., Duhrsen, U. & Moroy, T. (2002) *Nat Genet* 28, 28.

15. Saleque, S., Cameron, S. & Orkin, S. H. (2002) *Genes Dev* 16, 301-6.

16. Hock, H., Hamblen, M. J., Rooke, H. M., Traver, D., Bronson, R. T., Cameron, S. & Orkin, S. H. (2003) *Immunity* 18, 109-20.

17. Manzanares, M., Locascio, A. & Nieto, M. A. (2001) *Trends Genet* 17, 178-81.

18. Mayor, R., Morgan, R. & Sargent, M. G. (1995) *Development* 121, 767-77.

19. Inukai, T., Inoue, A., Kurosawa, H., Goi, K., Shinjyo, T., Ozawa, K., Mao, M., Inaba, T. & Look, A. T. (1999) *Mol Cell* 4, 343-52.

20. Cano, A., Perez-Moreno, M. A., Rodrigo, I., Locascio, A., Blanco, M. J., del Barrio, M. G., Portillo, F. & Nieto, M. A. (2000) *Nat Cell Biol* 2, 76-83.

21. Batlle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J. & Garcia De Herreros, A. (2000) *Nat Cell Biol* 2, 84-9.

22. Goyal, R. K., Lin, P., Kanungo, J., Payne, A. S., Muslin, A. J. & Longmore, G. D. (1999) *Mol Cell Biol* 19, 4379-89.

23. Kanungo, J., Pratt, S. J., Marie, H. & Longmore, G. D. (2000) *Mol Biol Cell* 11, 3299-313.

24. Kiss, H., Kedra, D., Yang, Y., Kost-Alimova, M., Kiss, C., O'Brien, K. P., Fransson, I., Klein, G., Imreh, S. & Dumanski, J. P. (1999) *Hum Genet* 105, 552-9.

25. Ayyanathan, K., Lechner, M. S., Bell, P., Maul, G. G., Schultz, D. C., Yamada, Y., Tanaka, K., Torigoe, K. & Rauscher, F. J., 3rd (2003) *Genes Dev* In Press.

26. Ayyanathan, K., Fredericks, W. J., Berking, C., Herlyn, M., Balakrishnan, C., Gunther, E. & Rauscher, F. J., 3rd (2000) *Cancer Res* 60, 5803-14.

27. Ryan, R. F., Schultz, D. C., Ayyanathan, K., Singh, P. B., Friedman, J. R., Fredericks, W. J. & Rauscher, F. J., 3rd (1999) *Mol Cell Biol* 19, 4366-78.

28. Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J., 3rd (2002) *Genes Dev* 16, 919-32.

29. Freyd, G., Kim, S. K. & Horvitz, H. R. (1990) *Nature* 344, 876-9.

30. Nix, D. A. & Beckerle, M. C. (1997) *J Cell Biol* 138, 1139-47.

31. Wadman, I. A., Osada, H., Grutz, G. G., Agulnick, A. D., Westphal, H., Forster, A. & Rabbitts, T. H. (1997) *Embo J* 16, 3145-57.

32. Visvader, J. E., Mao, X., Fujiwara, Y., Hahm, K. & Orkin, S. H. (1997) *Proc Natl Acad Sci U S A* 94, 13707-12.

33. McLoughlin, P., Ehler, E., Carlile, G., Licht, J. D. & Schafer, B. W. (2002) *J Biol Chem* 277, 37045-53.

34. Morlon, A. & Sassone-Corsi, P. (2003) *Proc Natl Acad Sci U S A* 100, 3977-82.

35. Sugihara, T. M., Bach, I., Kioussi, C., Rosenfeld, M. G. & Andersen, B. (1998) *Proc Natl Acad Sci U S A* 95, 15418-23.

36. Howard, P. W. & Maurer, R. A. (2001) *J Biol Chem* 276, 19020-6.

37. Chang, D. F., Belaguli, N. S., Iyer, D., Roberts, W. B., Wu, S. P., Dong, X. R., Marx, J. G., Moore, M. S., Beckerle, M. C., Majesky, M. W. & Schwartz, R. J. (2003) *Dev Cell* 4, 107-18.

38. Muller, J. M., Metzger, E., Greschik, H., Bosserhoff, A. K., Mercep, L., Buettner, R. & Schule, R. (2002) *Embo J* 21, 736-48.

39. Ng, E. K., Chan, K. K., Wong, C. H., Tsui, S. K., Ngai, S. M., Lee, S. M., Kotaka, M., Lee, C. Y., Waye, M. M. & Fung, K. P. (2002) *J Cell Biochem* 84, 556-66.

40. Taniguchi, Y., Furukawa, T., Tun, T., Han, H. & Honjo, T. (1998) *Mol Cell Biol* 18, 644-54.

41. Brown, S., McGrath, M. J., Ooms, L. M., Gurung, R., Maimone, M. M. & Mitchell, C. A. (1999) *J Biol Chem* 274, 27083-91.

42. Degenhardt, Y. Y. & Silverstein, S. (2001) *J Virol* 75, 11791-802.

43. Petit, M. M., Fradelizi, J., Golsteyn, R. M., Ayoubi, T. A., Menichi, B., Louvard, D., Van de Ven, W. J. & Friederich, E. (2000) *Mol Biol Cell* 11, 117-29.

44. Bach, I., Rodriguez-Esteban, C., Carriere, C., Bhushan, A., Krones, A., Rose, D. W., Glass, C. K., Andersen, B., Izpisua Belmonte, J. C. & Rosenfeld, M. G. (1999) *Nat Genet* 22, 394-9.

Figure Legends:

Figure 1. Characteristics of SNAG-zinc finger transcription factors. **A.** Alignment of human SNAG-ZFPs. Presence of a highly conserved NH_2-terminal SNAG repression module and a variable COOH-terminal zinc finger arrays are highlighted. The Slug and Scratch ZFPs also contain conserved slug and scratch domains, respectively. **B.** SNAG domain homology. The SNAG repression domain extends to ~20 amino acids, of which the NH_2-terminal first seven amino acids are strictly conserved. **C.** Diagrammatic representation of the wt and mutant SNAG-PAX3 repressors. **D.** Protein expression. The expression plasmids mentioned in panel **C** were transfected into COS-1 cells, ^{35}S-labeled cell extracts were immunoprecipitated with α-PAX3 IgG, and analyzed by SDS-PAGE and fluorography. **E.** Gel Shift. Nuclear extracts made from transfected COS-1 cells were incubated with ^{32}P-labeled e5 probe and electrophoresed on a non-denaturing PAGE. The arrow denotes the shifted bands. **F.** Transcriptional repression competence of SNAG-PAX3 repressors. The wt and mutant SNAG-PAX3 expression plasmids were transfected into NIH3T3 cells along with PAX3-luciferase reporter and CMV-lacZ plasmids. Normalized luciferase activity obtained for each construct is depicted.

Figure 2. Identification and characterization of SNAG-Associated Proteins (SNAPs). **A.** Characteristics of SNAG-LEXA repressors. The wt & mutant SNAG-LEXA repressors were tested for stable protein expression, and also monitored for repression of a LEXA-luciferase reporter. **B&C.** Specificity of SNAPs. **B.** The SNAP13 and SNAP20 clones bind to wt SNAG, but not to SNAG (FLV-AAA), or unrelated baits in the mating assay. **C.** These SNAPs bind to SNAG domains that retain repression activity, but not to either LEXA-DBD, or to mutants that abolish repression. Note that the KK-AA mutant is not stably expressed either in COS-1 (Panel **A**) or in yeast (data not shown), which accounts for its poor association. **D.** Group 3 LIM domain proteins. The SNAG-interacting regions of SNAP13 and SNAP20 contain two NH_2-terminal LIM domains of LIMD1 and Ajuba, respectively. The characteristic gly-, pro-rich regions and nuclear export signals (NES) present in these family members are highlighted. **E.** The signature Cys-His spacing ($CX_2CX_{16-23}HX_2CX_2CX_2CX_{16-21}CX_{2-3}C/H/D$ where C is cysteine, H is histidine, D is aspartic acid, and X is any amino acid), and conserved residues (highlighted in black bars) characteristic to LIM domains are shown as exemplified for by Ajuba LIM domains. **F.** Phylogenetic tree analysis. Sequence conservation between the LIM domains of the indicated three LIM domain proteins was determined using DNASTAR software.

Figure 3. Ajuba functions as a SNAG corepressor. **A.** Ajuba interacts with repression competent SNAG domains *in vivo*. The wt & mutant SNAG-LEXA plasmids were either transfected alone or co-transfected along with Ajuba into COS-1 cells. The ^{35}S-labeled cell extracts were immunoprecipitated with α-LEXA or α-MYC antibodies and analyzed. Arrow indicates the presence of co-precipitated Ajuba protein. **B.** Enhancement of SNAG mediated transcriptional repression by Ajuba. The wt & mutant SNAG-PAX3 or SNAG-

LEXA plasmids were either transfected alone or co-transfected with MYC-Ajuba. Relative luciferase activities obtained after normalization are depicted.

Figure 4. SNAG domain co-localizes with Ajuba in a repression-competence dependent manner. **A.** Localization patterns of Ajuba and SNAG-LEXA proteins. The MYC-Ajuba, wt, and mutant SNAG-LEXA plasmids were transfected individually into NIH3T3 cells and immunostaining was carried out with α-LEXA or α-MYC antibodies. Note that the Ajuba remains mainly cytoplasmic while the wt and mutant SNAG-LEXA proteins are predominantly nuclear. **B.** Co-localization of repression competent SNAG-LEXA proteins with MYC-Ajuba protein. The wt & mutant SNAG-LEXA plasmids were co-transfected along with the MYC-Ajuba plasmid into NIH3T3 cells. "Binding," indicates interaction in yeast two-hybrid analysis. "Repression" refers to the transcriptional competence in reporter assays. A representative cell is shown for each transfected population.

Figure 5. SNAG-mediated transcriptional repression mechanisms on a chromatinized locus. **A**. Development of a dual plasmid system. The SNAG-PAX3-HBD chimeric repressor protein is constitutively expressed from the CMV promoter. The reporter plasmid contains six repeats of PAX3 DNA binding motifs and a HSV thymidine kinase (TK) promoter that controls the basal expression of the luciferase gene. This plasmid also contains a zeocin^R cassette for drug selection. The primer-pair PBS1&PBS2 amplify the PAX3 binding sites (344 bp) while the TKP1&LUC1 primer-pair amplify the TK promoter region (257 bp). **B,C&D**. Chromatin IPs identifies putative components of the SNAG repression pathway. **B.** 4-OHT dependent enrichment of the SPHBD at the PAX3 sites. Chromatin prepared from – or + 4-OHT treated SPHL11 and SPHL20 cells were

immunoprecipitated with α-PAX3 (10 μg) and analyzed by quantitative PCR using PBS1&PBS2 primer-pair. Arrow indicates the PAX3-site fragment (344 bp) amplified. **C**. Ajuba and LIMD1 proteins were inducibly recruited to the TK promoter. ChIP experiments were done with - or + 4-OHT treated SPHL11 cells using α-Ajuba (20 μl) and α-LIMD1 (20 μl), α-MeK9 (5 μl) antisera. The immunoprecipitated DNA was amplified using TKP1&LUC1 primer-pair. Arrow denotes the amplified TK promoter fragment (257 bp). **D**. 4-OHT dependent promoter hypoacetylation. ChIP experiments were done with - or + 4-OHT treated SPHL11 & SPHL 20 cell lines using α–AcH3 IgG (5 μl) or α–AcH4 (5 μl) IgG and the retained DNA was tested by quantitative PCR using TKP1&LUC1 primer-pair. Arrow indicates the amplified TK promoter fragment (257 bp).

The LIM protein AJUBA recruits protein arginine methyltransferase 5 (PRMT5) to mediate SNAIL-dependent transcriptional repression

Zhaoyuan Hou[1], Hongzhuang Peng[1], Kasirajan Ayyanathan[3], Kaiping Yan[1], Ellen Langer[2], Gregory D. Longmore[2], Frank J. Rauscher III [1*]

1. The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104
2. Department of Medicine, Washington University, St. Louis, MO 63110, USA
3. Department of Biological Sciences, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431

*Corresponding Author : The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104. E-mail: Rauscher@wistar.org; Tel : (215) 898 0995; FAX: (215) 898 3929

Running title: PRMT5 mediates SNAIL-dependent transcriptional repression

Abbreviations: PRMT5, protein arginine methyltransferase 5; EMT, epithelial-mesenchymal transition. ChIP, Chromatin immunoprecipitation

Word count:

Abstract: 176

Materials and methods: 1060

Abstract

The SNAIL transcription factor contains C-terminal tandem zinc finger motifs and an N-terminal SNAG repression domain. The SNAIL family has recently emerged as major contributors to the processes of development and metastasis via regulating epithelial-mesenchymal differentiation events during embryonic development and tumor progression. However, the mechanisms by which SNAIL represses gene expression are largely undefined. Previously we demonstrated that AJUBA, a LIM domain protein, functions as a co-repressor for SNAIL and may serve as a platform for assembly of chromatin modifying factors. Here, we describe the identification of protein arginine methyltransferase 5 (PRMT5) as an effector recruited to SNAIL through direct binding to AJUBA which functions to repress the SNAIL target gene, *E-cadherin*. PRMT5 binds to the non-LIM region of AJUBA, and is translocated into the nucleus in a SNAIL and AJUBA-dependent manner. Depletion of PRMT5 in p19 cells stimulates *E-cadherin* expression, and the ternary complex can be found at the the proximal promoter region of the *E-cadherin* gene, concomitant with increased arginine methylation of histones at the locus. Together, these data suggest that PRMT5 is an effector of SNAIL-dependent gene repression.

Introduction

The SNAG family of zinc finger transcription factors in vertebrates include GFI-1A, GFI-1B, the insulinoma-associated protein IA-1, the homeobox protein GSH-1, and the SNAIL/SLUG family. These proteins play important roles in the regulation of development, stem cell self-renewal and tumor progression (4, 19, 44). They share a common set of functional domains: a C-terminal DNA binding domain composed of 5 –7 regularly spaced, Cys2-His2 zinc fingers and a highly conserved N-terminal repression domain designated SNAG. The SNAG motif was first identified from GFI-1 protein and comprises the first 21 amino acid residues in the N-terminus. The SNAG domain is a potent and transferable repression motif (19, 44). However, unlike other repression domains which are associated with zinc finger proteins such as the KRAB domain and the BTB-POZ domain whose mechanisms of repression are well established, little is known about the mechanisms of the SNAG domain-mediated repression (8, 14).

The SNAIL protein has emerged as a potent regulator to the processes of embryonic development and tumor progression by regulating the epithelial-mesenchymal transition (EMT) (4, 31). In mammalian cells, SNAIL induces EMT at least partially due to repression of the *E-cadherin* gene, thereby altering cell adhesion (5). The SNAIL protein has been found in multiprotein complexes containing histone deacetylases, mSIN3A and LOXL2/3 (34, 35). However, the biological significance of these interactions and how SNAIL mediates functional protein complex assembly at specific promoters in the context of chromatin remains undefined.

We previously identified novel co-repressors directly bound to the SNAG domains of GFI1A and SNAIL using yeast two-hybrid assays. The AJUBA protein was

identified as a prospective candidate which binds to the minimal SNAG domain. AJUBA is a multiple LIM domain-containing protein, and belongs to the ZYXIN family of LIM proteins (18). This family includes AJUBA, LIMD1, WTIP, ZYXIN, LPP, and TRIP6. The AJUBA/ZYXIN family is characterized by three tandem C-terminal LIM domains and unique N-terminal regions designated the preLIM regions (18, 23). The AJUBA protein is predominantly cytoplasmic, but can shuttle between the nucleus and cytoplasm (24). The AJUBA protein may function as a scaffold protein to assemble multiple cytoplasmic protein complexes involved in the processes of cell adhesion, migration, mitosis and cell differentiation (13, 18, 20). However, its role in the nucleus as a regulator of gene expression is poorly defined.

Detailed characterization of the interaction between AJUBA and SNAIL by *in vivo* and *in vitro* approaches demonstrated that AJUBA functions as a SNAIL co-repressor to repress the *E-cadherin* gene and is recruited to the endogenous *E-cadherin* promoter in a SNAIL-dependent manner (Langer et al, manuscript submitted). Expression of *Ajuba* during *Xenopus* development parallels that of *Snail*, and *Ajuba* cooperates with *Snail* during *Xenopus* neural crest development. Since AJUBA itself does not contain apparent enzymatic activity, we postulated that AJUBA may recruit other effectors to the SNAG domain of SNAIL to modify chromatin structure.

In the present study, we purified AJUBA interacting proteins and describe the protein arginine methyltransferase 5 (PRMT5) as a candidate. PRMT5 is a type II protein arginine methyltransferase and plays important role in the regulation of gene transcription (26). The studies provide strong evidence that PRMT5 is a key component of the SNAIL silencing complex through binding to AJUBA.

Materials and methods

Plasmids. The Myc epitope-tagged *Ajuba* plasmids pMEX-*myc*-*Ajuba* have been described (18). All *Ajuba* mutants and truncations were made using the QuikChange Site-Directed Mutagenesis procedures following the manufacturer's protocol (Stratagene, La Jolla, CA), and all mutants were confirmed by DNA sequencing. The *Ajuba* cDNAs were subcloned from pMEX-Myc-*Ajuba* via digestion with *Bam* HI and *Xho* I and insertion into the pcDNA3.1-N-Flag vector to create a Flag epitope-tagged AJUBA fusion protein in the N-terminus. The pcDNA-RFP-*Ajuba* and pGL2-*E-cadherin* luciferase reporter plasmids have been previously described(37, 40). The pcDNA-Flag-*PRMT5* plasmids was provided by Dr. Dreyfuss (15). The Sport6-CMV-*Snail* (murine) plasmid was purchased from Open Biosystems (Huntsville, AL).

The GST-AJUBA(244-350aa) and leucine to alanine substitutions were generated by PCR amplification of the DNA fragment encoding aa residues 244-350 of murine AJUBA. The PCR products were cloned into the *Bam* HI and *Eco* RI sites of the pGEX-4T-1. The truncated PRMT5 plasmids containing PRMT5(1-170), PRMT5(169-422), and PRMT5(421-637) were generated by amplifying the DNA fragment encoding the indicated residues of human PRMT5 and the resulting fragments were cloned into the pET-28a vector.

Cell culture, transfections and luciferease reporter assays. HEK293 cells, U2OS cells and p19 cells were maintained in DMEM containing 10% FBS, 2 mM L-glutamine, and penicillin (50 U/ml)/streptomycin (50 μg/ml) at 37 C under 5% CO_2 in a humidified chamber.

For transfection, HEK293 cells were seeded at 5 x 10^4 cells per well in 24-well plates. The *ß-galactosidase* plasmid (50 ng) and pGL2-E-cad-Luc reporter (200 ng) along with *Snail* and/or *Ajuba* encoding plasmids were transiently transfected into the cells with Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA) following the manufacturer's protocol. Twenty four hours post transfection, cells were harvested and lysed. Luciferase and ß-galactosidase activities were measured with the luciferase reporter gene assay kit (Promega, Madison, WI) and the chemiluminescent reporter assay for ß-galactosidase kit (Clontech, Mountain View, CA), respectively. The transfection efficiency among plates was normalized to ß-galactosidase activity and all transfections were repeated three times in duplicate.

Affinity purification of a native AJUBA complex. To purify AJUBA-associated proteins, a Flag-tagged, full-length *Ajuba* cDNA in the pcDNA3.1 vector was stably expressed in HEK293 cells. Single cell clones were selected with G418 and screened by western blot using anti-Flag antibody. A cell clone expressing the Flag-AJUBA protein at level comparable to that of the endogenous AJUBA was chosen for the purification. A total of 5x10^9 cells were lysed in buffer A containing 20mM Tris HCl (pH 8.0), 150mM NaCl, 2.5mM EDTA, 0.5% NP40, 0.2 mM PMSF, and 0.5 mM DTT. Cell lysates were pre-cleared with the protein A agarose beads for 2 hours, and then incubated with the anti-Flag agarose M2 beads (Sigma, St. Louis, MO) at 0.5 ml of beads per 100 mg of cell lysate for 2 hours to overnight with rotation. The M2 beads were washed 4 times with buffer BC500 containing 20 mM Tris-HCl (pH 7.8), 500 mM KCl, 0.2 mM EDTA, 10% glycerol, 10 mM β-mercaptoethanol, 0.2% NP40, 0.2 mM

PMSF and 1μg/ml of aprotinin, leupeptin, and pepstatin. The protein complex was eluted with the Flag-peptides (Sigma) at 100μg/ml in buffer BC100 containing 20 mM Tris-HCl (pH 7.8), 50 mM KCl, 0.2 mM EDTA, 10% glycerol, 10 mM β-mercaptoethanol, 0.2 mM PMSF and 1μg/ml of aprotinin, leupeptin, and pepstatin. The eluted proteins were resolved on 4-12% SDS-PAGE gels for western blot, silver and colloidal staining analyses. The proteins were excised from the gel and identified by standard mass spectrometry at the Wistar Institute Cancer Center, Proteomics Core Facility.

Co-immunoprecipiation, western blot, immunofluorescence and antibodies. Myc-*Ajuba*, Flag-*PRMT5* and/or Flag-*Snail* plasmids were transiently transfected into HEK293 cells and 24 hours post transfection, cells were lysed in buffer A. Co-immunoprecipiations were performed with either anti-Myc or anti-Flag reagents. The western blots and immunofluorescences were previously described (21, 41). Mouse monoclonal anti-Myc (Invitrogen), anti-Flag (Sigma), Rabbit polyclonal anti-SNAIL (Santa Cruz Biotechnology, Sant Cruz, CA), anti-H4R3 (UPSTATE, Charlottesville, VA), anti-E-cadherin (Cell Signaling, Danvers, MA) antibodies were purchased. The rabbit polyclonal anti-AJUBA antibody was raised by immunizing rabbits with a bacterial expressed 6-His fusion protein of murine AJUBA (amino acids 1-216) as the antigen.

SiRNA knockdown, the methyltransferase inhibitor 5'-Deoxy-5'-methyl-thioadenosine (MTA) treatment and RT-PCR. Smart pool siRNAs targeting murine *Ajuba* and *PRMT5* (Dharmacon, Lafayette, CO) were transfected into the cells with the Lipofectamine 2000 reagent (Invitrogen). 5'-Deoxy-5'-methyl-thioadenosine (Sigma)

was disolved in DMSO. p19 cells were seeded at $1.5x10^5$ cells/well in 6-well cell culture plates on day 0. MTA was added into the media on day 1 at concentrations of 100uM and 200uM for 48 hours.

Total RNA from p19 cells was isolated with the RNeasy Kit (Qiagen, Valencia, CA). RNA was treated with RQ DNase I to remove any genomic DNA contamination. Two micrograms of the treated total RNA were used for cDNA synthesis in a 20 μl reaction with the Superscript II reverse transcriptase (Invitrogen). The primer pairs used for *E-cadherin* and *GAPDH* amplification were: sense 5'-GAGAACGGTGGTCAAAGAGC-3', anti-sense 5'-CATCTCCCATGGTGCCACAC-3, and sense 5' ACCACAGTCCATGCCATCAC, anti-sense: TCCACCCCCTGTTGCTGTA respectively. PCR amplification was carried out using *Taq* DNA Polymerase (Promega) at 94 C for 15 sec, at 60 C for 15 sec, and 72 C for 60 sec.

Chromatin immunoprecipitation. Chromatin immunoprecipitation (ChIP) experiments were carried out in HEK293 cells stably expressing the Flag-*Snail* cDNA. HEK 293-Flag-*Snail* and HEK 293-vector cell lines were established by transfection of pcDNA3.1 Flag-*Snail* and parental vectors into HEK293 cells, and selected with zeocin at 400ng/ml. The expression of the Flag-SNAIL protein was confirmed by western blot. To prepare cells for ChIP, HEK293-Flag-*Snail* and HEK293-vector cells were grown in 150 mm plates to 70-90% confluency, and fixed by addition of 574 μl of 37% formaldehyde directly into 20 ml of growth medium to a final concentration of 1% for 20 minutes in the cell culture incubator. The crosslinking reaction was stopped by addition

of 1.25 ml of 2M glycine in PBS buffer at room temperature for 5 minutes. Cells were harvested and the ChIP assays were performed according to the protocol supplied with the EZ-CHIP kit (UPSTATE, Cat. 17-371). The immunoprecipitated DNAs were amplified with the primer set (Primer 1) P1 5'-AATCAGAACCGTGCAGGTCC-3' and P2 5'-ACAGGTGCTTTGCAGTTCCG-3'. This 250bp amplicon flanks the three E-boxes located in the proximal promoter region of the *E-cadherin* gene. A second set of primers (Primer 2), P3 5'-GGCTCAAGCTATCCTTGCAC-3', and P4 5'-GTGCAGTGGCTCAT-GTCTGT-3' was used to amplify a 197bp fragment, encoded by exon 16 of the *E-cadherin* gene. The PCR fragments were cloned and their identities were confirmed by DNA sequencing. For quantitation, the PCR products were resolved on 2% agarose gels and visualized with ethidium bromide.

Results

PRMT5 was identified as one of the AJUBA-interacting proteins

To isolate potential AJUBA-interacting proteins which may repress the *Snail* target genes such as *E-cadherin*, we chose a cell system which supports SNAIL-mediated represssion. Since HEK293 cells have been widely used for purification of protein complexes, we first tested whether SNAIL can repress *E-cadherin* expression in this cell lines. The E-cadherin-Luc reporter contains three SNAIL binding sites, which are located in the proximal promoter region of the *E-cadherin* gene, a promoter which is responsive to SNAIL in a variety of cell types (5, 34, 39). HEK293 cells were transiently transfected with *Snail* and *Ajuba* encoding plasmids alone or in combination along with the E-cadherin-luc reporter. *Snail* by itself was able to repress the

transcription of the *E-cadherin* promoter-driven luciferase reporter in a dosage dependent manner (Figure 1B, top panel). When *Ajuba* and *Snail* were co-expressed in HEK293 cells, we observed increased repression over *Snail* alone (Figure 1B, low panel). These results demonstrate that HEK293 cells contain the factors required for SNAIL and AJUBA to repress *E-cadherin*. To purify these factors, we selected a cell clone stably expressing the Flag-AJUBA protein comparable to that of the endogenous level and performed affinity purification (Figure 1C). As expected, the major protein eluted in the affinity purification migrated at MW 80kDa was the Flag-AJUBA protein. Other proteins migrated at MW 95kDa, 70kDa, and 50 kDa were observed in multiple independent purification experiments. The major protein at MW 95kDa was identified as PRMT5 and was pursued as a candidate AJUBA-interacting protein. The MEP50 (MW 50kDa) protein, a known co-factor for PRMT5, was also observed in the purification (16, 17).

To confirm the interaction between AJUBA and PRMT5, we first transiently co-transfected full-length Myc-*Ajuba* and Flag-*PRMT5* cDNAs into HEK293 cells. Reciprical co-immunoprecipitations showed that PRMT5 and AJUBA can interact with each other in HEK293 cells (Figure 2A). To confirm that an AJUBA-PRMT5 interaction occurs with the endogenous proteins, whole cell lysates from p19 cells were used for immunoprecipitation. Indeed, the endogenous AJUBA was able to immunoprecipitate the endogenous PRMT5 (Figure 2B). Collectively, these data suggest that PRMT5 is a novel AJUBA-interacting protein.

The leucine rich motif in the preLIM region of AJUBA is essential for PRMT5 binding

The aa sequence of the AJUBA protein predicts two structurally distinct regions: the preLIM region in the N-terminus and the LIM domains in the C-terminus (Figure 3A). To determine which regions in AJUBA are responsible for the PRMT5 binding, plasmids encoding Myc-tagged, full-length AJUBA and the two truncated forms, preLIM and LIM, were co-transfected with the Flag-PRMT5 plasmid into HEK293 cells. The immunoprecipitation experiments showed that the preLIM region bound PRMT5, while the LIM domain did not (Figure 3B, Lanes 2, 4, 6). To further identify which specific motifs or residues in the preLIM region are critical for the interaction between AJUBA and PRMT5, we made progressive C-terminal deletions in the preLIM region (Figure 3A). These experiments confirmed that the preLIM bound to PRMT5. Interestingly, the truncated preLIM-312 protein showed stronger binding than that of the complete preLIM region, suggesting that the region encompaning aa 312-347 may contain inhibitory motifs to prohibit PRMT5 binding (Figure 3C, Lanes 2-3). The proteins preLIM-279, 255 and 210 failed to bind PRMT5. These data indicate that the region between residues 279 and 312 of AJUBA is essential for PRMT5 binding. Analysis of this 33 aa region (279-312) revealed a 13 residue leucine-rich region (DELTALLRLTVAT) with potential α helical character (Figure 4A). This region is also a potential nuclear export signal (18). To determine the role of these leucines in PRMT5 binding, we mutated them individually or in combination to alanine. Mutation of individual leucine residues did not significantly affect the interaction, while simultaneous mutation of leucines 292, 293 and 295 (AJUBA L3A) greatly attenuated PRMT5 binding (Figure 4B).

To identify the regions in the PRMT5 protein that can interact with AJUBA, we used the full-length PRMT5 and three truncations (Figure 4C). These fragments were translated *in vitro* and used for *in vitro* GST-AJUBA (244-350aa) binding assays. The full-length PRMT5 and the three truncations bound to GST-AJUBA (wild-type), while no binding was observed for GST-AJUBA L3A mutant (data not shown). These data suggest that multiple domains in the PRMT5 protein can interact with AJUBA.

Co-localization of AJUBA and PRMT5 in U2OS cells

To detect the subcellular localization of AJUBA and PRMT5, we performed immunofluorescent assays in U2OS cells. Plasmids encoding Myc-AJUBA and Flag-PRMT5 were transiently transfected into U2OS cells. When expressed alone, both the AJUBA and PRMT5 proteins were predominantly cytoplasmic with similar distribution patterns (Figure 5 A, C). When co-expressed, the distribution of each protein was not significantly changed (Figure 5 D). Similar to the wild type AJUBA, the AJUBA L3A mutant was also localized in the cytoplasm and mutation of the three leucines to alanines did not apparently affect its localization (Figure 5 B,E).

SNAIL, AJUBA and PRMT5 are found in the same complex

To determine whether the AJUBA protein functions as an adaptor to bridge SNAIL and PRMT5 proteins, we looked for ternary complex formation using co-immunoprecipitation. We transfected the Myc-*Ajuba* encoding plasmid, together with Flag-*Snail* and Flag-*PRMT5* encoding plasmids into HEK293 cells. We observed that AJUBA could simultaneously associate with PRMT5 and SNAIL (Figure 6A, lane 5).

This was surprising given that the AJUBA and PRMT5 proteins are predominately cytoplasmic, whereas SNAIL is mostly nuclear. We speculated that PRMT5 can be translocated into the nucleus by forming a complex with SNAIL and AJUBA. To visualize such an interaction, we transfected *Snail*, Flag-*PRMT5*, RFP-*Ajuba* and/or pcDNA-*RFP* encoding plasmids into U2OS cells. When expressed alone, the SNAIL protein was localized in the nucleus, while the AJUBA and PRMT5 proteins were in the cytoplasm. The RFP alone was found in both the cytoplasm and the nucleus and the RFP-AJUBA fusion protein showed identical localization compared to the wild type AJUBA (Figure 6B). However, co-expression of AJUBA and SNAIL affected both of their localization: the SNAIL protein was retained in the cytoplasm, and concomitantly a significant amount of the AJUBA protein was localized in the nucleus (Figure 6D). In contrast, co-expression of PRMT5 and SNAIL revealed no apparent effect on either localization (Figure 6E). Strikingly, when AJUBA, SNAIL, and PRMT5 were co-expressed, a significant amount of the PRMT5 protein was relocated to the nucleus and was co-localized with nuclear SNAIL and AJUBA (Figure 6F). Taken together, the results suggest that SNAIL, AJUBA and PRMT5 can form a complex in the nucleus.

Modulation of AJUBA and PRMT5 in p19 cells results in upregulation of *E-cadherin* expression.

Since we showed evidence of nuclear localized SNAIL-AJUBA-PRMT5 complex, we sought to determine whether the well-established SNAIL target gene *E-cadherin* was targeted by this ternary complex. It has been previously found that induction of *Snail* in p19 cells repressed *E-cadherin* gene expression, while depletion of

Snail resulted in upregulation of the *Snail*-target genes (9, 30). An siRNA pool targeting murine *Ajuba* was transfected into the p19 cells. The level of the AJUBA protein was significantly decreased by the siRNA (Figure 7A). Moreover, the expression of *E-cadherin* at both the mRNA and protein levels was significantly increased (Figure 7A, B). These results suggest that AJUBA is involved in the repression of *E-cadherin* gene expression.

To determine the role of PRMT5 in the regulation of the *E-cadherin* gene expression, we treated p19 cells with the methyltransferase inhibitor MTA, which has been shown to block PRMT5 function (38). Treatment of the p19 cells with MTA at the doses of 100uM and 200uM stimulated *E-cadherin* expression compared to the DMSO control (Figure 7C). As a more specific approach, we employed siRNA of knockdown PRMT5. The siRNA pool targeting *PRMT5* was transfected into p19 cells and the protein levels of PRMT5 was significantly decreased (Figure 7D). Moreover, *E-cadherin* expression was increased in PRMT5 knockdown cells as shown by western blot analyses. These data strongly suggest that both AJUBA and PRMT5 mediate SNAIL-dependent repression of *E-cadherin* and that the methyltransferase activity of PRMT5 is required.

SNAIL, AJUBA and PRMT5 bind to the endogenous promoter of *E-cadherin*

To examine whether SNAIL, AJUBA and PRMT5 form a functional multi-protein complex to repress the SNAIL target gene in living cells, we employed ChIP assays to examine their association with the endogenous promoter of the *E-cadherin* gene.

Clonal HEK293-Flag-*Snail* cells were established by stable transfection and protein expression was confirmed by western blot (Figure 8B). Expression of exogenous Flag-SNAIL induced morphological changes in HEK293 cells as has been observed on other cell types (data not shown) (1, 7, 22, 43). Downregulation of the *E-cadherin* gene expression was observed in Flag-Snail transfected HEK293 cells (Figure 8B). The immunoprecipitated DNA fragments were examined by PCR amplification using a primer set (Primer1) which flanks the three SNAIL-binding sites located in the proximal promoter region of the *E-cadherin* gene. A second primer set (Primer 2) amplified a fragment which is present in exon 16 of the *E-cadherin* gene (Figure 8A). We observed that the proximal promoter of the *E-cadherin* gene flanking the three SNAIL binding sites was highly enriched by antibodies to SNAIL, AJUBA, and PRMT5 in HEK293-Flag-*Snail* cells. However, in the HEK293-*vector* cells this enrichment was not observed (Figure 8C). Using an antibody which detects methylated histone H4 at arginine 3 (H4R3) we detected increased H4 methylation at the promoter of the *E-cadherin* gene (Figure 8C). The fragment resided in exon 16 was not enriched by any of these antibodies. Together, these data suggest that the association of SNAIL, AJUBA and PRMT5 with the *E-cadherin* gene occurs at the proximal promoter of the *E-cadherin* gene.

Discussion

Herein, we describe the identification of PRMT5 as a repressor recruited to the SNAIL complex via interacting with the AJUBA co-repressor. We demonstrated that PRMT5 can form multi-protein complexes containing SNAIL and AJUBA which

function to repress the canonical SNAIL target gene *E-cadherin*. To support this, we showed that treatment of the p19 cells with the methyltransferase inhibitor MTA or with siRNA targeting PRMT5 stimulate *E-cadherin* expression. Further, PRMT5 was shown to specifically bind to the proximal promoter of the *E-cadherin* gene and concomitantly, the methylation status of histones at this locus represented by H4R3 is increased in the presence of the SNAIL. Together, these studies suggest that PRMT5 is a key mediator for the regulation of the *E-cadherin* gene expression and the methyltransferase acivity of PRMT5 is involved in the transcriptional repression of the SNAIL complex.

The PRMT5 protein is a member of type II protein arginine methyltransferases and can methylate transcription factors and histones on specific arginine residues to regulate gene expression (2, 12, 32, 33). For examples, PRMT5 was found to interact with BRG1 and BRM, components of human SWI/SNF chromatin remodeling complex to methylate histones H2A, and H4 on arginine 3, and H3 on arginine 8. These activities of PRMT5 result in the repression of genes such as *ST7* and *NM23* and promotion of a tumorigenic state in NIH3T3 cell (32, 33). PRMT5 can interact with Blimp1, a zinc finger transcriptional repressor, and supresses *Dhx38* gene expression by methylation of the histones H2A and H4 on arginine 3 (2). The PRMT5 protein has also been found to be part of E2F complex in the *cyclin E1* promoter, correlating with repression of transcription of *cyclin E1* gene (12).

As described above, this is good evidence that PRMT5 is involved in the transcriptional repression. Paradoxically, the majority of the PRMT5 protein at steady state is found in the cytoplasm. How PRMT5 is translocated, retained and targeted to specific genes in the nucleus is not clear. We demonstrate that PRMT5 can be

translocated into the nucleus via forming a complex with AJUBA and SNAIL. This SNAIL may function as a nuclear anchor to target PRMT5 to its target genes.

The MEP50 protein is also found in the AJUBA-PRMT5 complex. The MEP50 protein contains WD motifs and is constitutively associated with the PRMT5 protein (16, 17). Several WD motif-containing proteins were recently shown to be essential for global histone methylation and the regulation of gene transcription (42). The MEP50 protein was shown to direct PRMT5 to specific histones and is indispensible for PRMT5-dependent histone modification (17). Further work will be necessary to establish the role of MEP50 in SNAIL-mediated gene repression and PRMT5 function.

The SNAIL family of protein play key roles in regulation of EMT events during the development and metastasis, and also serve as an early markers for the malignant phenotype and prognosis (3, 4, 6, 10, 27, 29). Recently, *Snail* was shown to be spontaneously up-regulated during the process of tumor recurrence in mice and high level of the *Snail* expression strongly predicts decreased relapse-free survival in women with breast cancer (29). These observations strongly implicate a critical role of *Snail* in the process of breast cancer recurrence. Therefore, identification of proteins involved in SNAIL-dependent repression will not only shed new light on understanding the mechanims of SNAIL in EMT and tumor recurrence, but also provide new targets for potential drug development and diagnostics.

Previous studies have demonstrated the association of SNAIL with potential co-regulators through its SNAG domain. These include histone deacetylases 1 and 2 (HDAC1 and HDAC2), and the co-repressor mSIN3A (34). HDACs are commonly found in large protein complexes *in vivo* both in the cytoplasm and in the nucleus and

may direct regulation of gene expression, the cell cycle, differentiation, and DNA repair. HDAC1 and HDAC2 have been shown to associate with SMRT, the CoREST complex, mSIN3, N-CoR, Mi-2/NuRD and play essential roles in gene silencing (11, 36). However, how SNAIL mediates the complex assembly remains elusive. Here, we showed that AJUBA recruits PRMT5 via its preLIM region to the SNAG domain of SNAIL. Since both HDAC1 and HDAC2, and AJUBA-PRMT5 interact with the SNAG domain of SNAIL, it will be interesting to examine the compementary roles of these enzymes in SNAIL-mediated repression.

It is a reasonable assumption that, in addition to PRMT5, other factors are recruited to the SNAIL complex via interacting with AJUBA. Like many other LIM proteins, AJUBA has been shown to function as a scaffold protein and can interact with a variety of proteins including transcription regulators, kinases, and cytoskeleton proteins (13, 18, 20, 23-25, 28, 37). How these interactions are spatially and temporally regulated should shed new light on the roles of these important and clinically relevent transcription factors.

Acknowledgments

We thank Dr. Gideon Dreyfuss for the gift of the PRMT5 plasmid and Dr. Jae-Gahb Park for the E-cadherin promoter reporter plasmid. We also thank Drs. David E. White, Dmitri Negorev, and Alexey V. Ivanov for helpful discussion. F.J.R. was supported by grants from the National Health Institute.

References

1. **Aguilera, A., L. S. Aroeira, M. Ramirez-Huesca, M. L. Perez-Lozano, A. Cirugeda, M. A. Bajo, G. Del Peso, A. Valenzuela-Fernandez, J. A. Sanchez-Tomero, M. Lopez-Cabrera, and R. Selgas.** 2005. Effects of rapamycin on the epithelial-to-mesenchymal transition of human peritoneal mesothelial cells. Int J Artif Organs **28:**164-9.
2. **Ancelin, K., U. C. Lange, P. Hajkova, R. Schneider, A. J. Bannister, T. Kouzarides, and M. A. Surani.** 2006. Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. Nat Cell Biol **8:**623-30.
3. **Bagnato, A., and L. Rosano.** 2007. Epithelial-mesenchymal transition in ovarian cancer progression: a crucial role for the endothelin axis. Cells Tissues Organs **185:**85-94.
4. **Barrallo-Gimeno, A., and M. A. Nieto.** 2005. The Snail genes as inducers of cell movement and survival: implications in development and cancer. Development **132:**3151-61.
5. **Cano, A., M. A. Perez-Moreno, I. Rodrigo, A. Locascio, M. J. Blanco, M. G. del Barrio, F. Portillo, and M. A. Nieto.** 2000. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat Cell Biol **2:**76-83.
6. **Castro Alves, C., E. Rosivatz, C. Schott, R. Hollweck, I. Becker, M. Sarbia, F. Carneiro, and K. F. Becker.** 2007. Slug is overexpressed in gastric carcinomas and may act synergistically with SIP1 and Snail in the down-regulation of E-cadherin. J Pathol **211:**507-15.

7. **Cicchini, C., D. Filippini, S. Coen, A. Marchetti, C. Cavallari, I. Laudadio, F. M. Spagnoli, T. Alonzi, and M. Tripodi.** 2006. Snail controls differentiation of hepatocytes by repressing HNF4alpha expression. J Cell Physiol **209:**230-8.

8. **Collins, T., J. R. Stone, and A. J. Williams.** 2001. All in the family: the BTB/POZ, KRAB, and SCAN domains. Mol Cell Biol **21:**3609-15.

9. **de Boer, T. P., T. A. van Veen, M. F. Bierhuizen, B. Kok, M. B. Rook, K. J. Boonen, M. A. Vos, P. A. Doevendans, J. M. de Bakker, and M. A. van der Heyden.** 2007. Connexin43 repression following epithelium-to-mesenchyme transition in embryonal carcinoma cells requires Snail1 transcription factor. Differentiation **75:**208-18.

10. **Ding, J. X., Y. J. Feng, L. Q. Yao, M. Yu, H. Y. Jin, and L. H. Yin.** 2006. The reinforcement of invasion in epithelial ovarian cancer cells by 17 beta-Estradiol is associated with up-regulation of Snail. Gynecol Oncol **103:**623-30.

11. **Ekwall, K.** 2005. Genome-wide analysis of HDAC function. Trends Genet **21:**608-15.

12. **Fabbrizio, E., S. El Messaoudi, J. Polanowska, C. Paul, J. R. Cook, J. H. Lee, V. Negre, M. Rousset, S. Pestka, A. Le Cam, and C. Sardet.** 2002. Negative regulation of transcription by the type II arginine methyltransferase PRMT5. EMBO Rep **3:**641-5.

13. **Feng, Y., and G. D. Longmore.** 2005. The LIM protein Ajuba influences interleukin-1-induced NF-kappaB activation by affecting the assembly and activity of the protein kinase Czeta/p62/TRAF6 signaling complex. Mol Cell Biol **25:**4010-22.

14. **Friedman, J. R., W. J. Fredericks, D. E. Jensen, D. W. Speicher, X. P. Huang, E. G. Neilson, and F. J. Rauscher, 3rd.** 1996. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev **10:**2067-78.

15. **Friesen, W. J., S. Paushkin, A. Wyce, S. Massenet, G. S. Pesiridis, G. Van Duyne, J. Rappsilber, M. Mann, and G. Dreyfuss.** 2001. The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. Mol Cell Biol **21:**8289-300.

16. **Friesen, W. J., A. Wyce, S. Paushkin, L. Abel, J. Rappsilber, M. Mann, and G. Dreyfuss.** 2002. A novel WD repeat protein component of the methylosome binds Sm proteins. J Biol Chem **277:**8243-7.

17. **Furuno, K., T. Masatsugu, M. Sonoda, T. Sasazuki, and K. Yamamoto.** 2006. Association of Polycomb group SUZ12 with WD-repeat protein MEP50 that binds to histone H2A selectively in vitro. Biochem Biophys Res Commun **345:**1051-8.

18. **Goyal, R. K., P. Lin, J. Kanungo, A. S. Payne, A. J. Muslin, and G. D. Longmore.** 1999. Ajuba, a novel LIM protein, interacts with Grb2, augments mitogen-activated protein kinase activity in fibroblasts, and promotes meiotic maturation of Xenopus oocytes in a Grb2- and Ras-dependent manner. Mol Cell Biol **19:**4379-89.

19. **Grimes, H. L., T. O. Chan, P. A. Zweidler-McKay, B. Tong, and P. N. Tsichlis.** 1996. The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. Mol Cell Biol **16:**6263-72.

20. **Hirota, T., N. Kunitoku, T. Sasayama, T. Marumoto, D. Zhang, M. Nitta, K. Hatakeyama, and H. Saya.** 2003. Aurora-A and an interacting activator, the LIM protein Ajuba, are required for mitotic commitment in human cells. Cell **114:**585-98.

21. **Hou, Z., S. Srivastava, M. J. Mistry, M. P. Herbst, J. P. Bailey, and N. D. Horseman.** 2003. Two tandemly linked interferon-gamma-activated sequence elements in the promoter of glycosylation-dependent cell adhesion molecule 1 gene synergistically respond to prolactin in mouse mammary epithelial cells. Mol Endocrinol **17:**1910-20.

22. **Julien, S., I. Puig, E. Caretti, J. Bonaventure, L. Nelles, F. van Roy, C. Dargemont, A. G. de Herreros, A. Bellacosa, and L. Larue.** 2007. Activation of NF-kappaB by Akt upregulates Snail expression and induces epithelium mesenchyme transition. Oncogene.

23. **Kadrmas, J. L., and M. C. Beckerle.** 2004. The LIM domain: from the cytoskeleton to the nucleus. Nat Rev Mol Cell Biol **5:**920-31.

24. **Kanungo, J., S. J. Pratt, H. Marie, and G. D. Longmore.** 2000. Ajuba, a cytosolic LIM protein, shuttles into the nucleus and affects embryonal cell proliferation and fate decisions. Mol Biol Cell **11:**3299-313.

25. **Kisseleva, M., Y. Feng, M. Ward, C. Song, R. A. Anderson, and G. D. Longmore.** 2005. The LIM protein Ajuba regulates phosphatidylinositol 4,5-bisphosphate levels in migrating cells through an interaction with and activation of PIPKI alpha. Mol Cell Biol **25:**3956-66.

26. **Krause, C. D., Z. H. Yang, Y. S. Kim, J. H. Lee, J. R. Cook, and S. Pestka.** 2007. Protein arginine methyltransferases: evolution and assessment of their pharmacological and therapeutic potential. Pharmacol Ther **113:**50-87.

27. **Kurrey, N. K., A. K, and S. A. Bapat.** 2005. Snail and Slug are major determinants of ovarian cancer invasiveness at the transcription level. Gynecol Oncol **97:**155-65.

28. **Marie, H., S. J. Pratt, M. Betson, H. Epple, J. T. Kittler, L. Meek, S. J. Moss, S. Troyanovsky, D. Attwell, G. D. Longmore, and V. M. Braga.** 2003. The LIM protein Ajuba is recruited to cadherin-dependent cell junctions through an association with alpha-catenin. J Biol Chem **278:**1220-8.

29. **Moody, S. E., D. Perez, T. C. Pan, C. J. Sarkisian, C. P. Portocarrero, C. J. Sterner, K. L. Notorfrancesco, R. D. Cardiff, and L. A. Chodosh.** 2005. The transcriptional repressor Snail promotes mammary tumor recurrence. Cancer Cell **8:**197-209.

30. **Nakakura, E. K., D. N. Watkins, V. Sriuranpong, M. W. Borges, B. D. Nelkin, and D. W. Ball.** 2001. Mammalian Scratch participates in neuronal differentiation in P19 embryonal carcinoma cells. Brain Res Mol Brain Res **95:**162-6.

31. **Nieto, M. A.** 2002. The snail superfamily of zinc-finger transcription factors. Nat Rev Mol Cell Biol **3:**155-66.

32. **Pal, S., S. N. Vishwanath, H. Erdjument-Bromage, P. Tempst, and S. Sif.** 2004. Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes. Mol Cell Biol **24:**9630-45.

33. **Pal, S., R. Yun, A. Datta, L. Lacomis, H. Erdjument-Bromage, J. Kumar, P. Tempst, and S. Sif.** 2003. mSin3A/histone deacetylase 2- and PRMT5-containing Brg1 complex is involved in transcriptional repression of the Myc target gene cad. Mol Cell Biol **23:**7475-87.

34. **Peinado, H., E. Ballestar, M. Esteller, and A. Cano.** 2004. Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex. Mol Cell Biol **24:**306-19.

35. **Peinado, H., M. Del Carmen Iglesias-de la Cruz, D. Olmeda, K. Csiszar, K. S. Fong, S. Vega, M. A. Nieto, A. Cano, and F. Portillo.** 2005. A molecular role for lysyl oxidase-like 2 enzyme in snail regulation and tumor progression. Embo J **24:**3446-58.

36. **Peterson, C. L.** 2002. HDAC's at work: everyone doing their part. Mol Cell **9:**921-2.

37. **Pratt, S. J., H. Epple, M. Ward, Y. Feng, V. M. Braga, and G. D. Longmore.** 2005. The LIM protein Ajuba influences p130Cas localization and Rac1 activity during cell migration. J Cell Biol **168:**813-24.

38. **Richard, S., M. Morel, and P. Cleroux.** 2005. Arginine methylation regulates IL-2 gene expression: a role for protein arginine methyltransferase 5 (PRMT5). Biochem J **388:**379-86.

39. **Shin, Y., I. J. Kim, H. C. Kang, J. H. Park, H. R. Park, H. W. Park, M. A. Park, J. S. Lee, K. A. Yoon, J. L. Ku, and J. G. Park.** 2004. The E-cadherin -347G->GA promoter polymorphism and its effect on transcriptional regulation. Carcinogenesis **25:**895-9.

40. **Shin, Y., I. J. Kim, H. C. Kang, J. H. Park, H. W. Park, S. G. Jang, M. R. Lee, S. Y. Jeong, H. J. Chang, J. L. Ku, and J. G. Park.** 2004. A functional polymorphism (-347 G-->GA) in the E-cadherin gene is associated with colorectal cancer. Carcinogenesis **25:**2173-6.

41. **White, D. E., D. Negorev, H. Peng, A. V. Ivanov, G. G. Maul, and F. J. Rauscher, 3rd.** 2006. KAP1, a novel substrate for PIKK family members, colocalizes with numerous damage response factors at DNA lesions. Cancer Res **66:**11594-9.

42. **Wysocka, J., T. Swigut, T. A. Milne, Y. Dou, X. Zhang, A. L. Burlingame, R. G. Roeder, A. H. Brivanlou, and C. D. Allis.** 2005. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell **121:**859-72.

43. **Zhou, B. P., J. Deng, W. Xia, J. Xu, Y. M. Li, M. Gunduz, and M. C. Hung.** 2004. Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. Nat Cell Biol **6:**931-40.

44. **Zweidler-Mckay, P. A., H. L. Grimes, M. M. Flubacher, and P. N. Tsichlis.** 1996. Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. Mol Cell Biol **16:**4024-34.

Figure 1. Affinity purification of AJUBA interacting proteins. (A) the diagram showing the architectures of the SNAIL and AJUBA proteins. (B) AJUBA can augment SNAIL-mediated repression on the promoter luciferase reporter activities of *E-cadherin* in HEK293 cells. (C) colloidal staining shows the potential AJUBA-interacting proteins

purified from HEK293 cells. * represents Actins. The pre-stained molecular weight marker (Benchmarker, Invitrogen) migrates 15-20kDa faster than predicted MW.

Figure 2. PRMT5 is a novel AJUBA-interacting protein. (A) the exogenously expressed PRMT5 and AJUBA proteins interact in HEK293 cells. (B) the endogenous AJUBA and PRMT5 proteins interact in p19 cells. Immunoprecipitation was performed with anti-AJUBA polyclonal antibody and pre-immune rabbit IgG was used as control. The immunoprecipitated proteins were resolved on SDS-PAGE gel and western blot was performed with monoclonal anti-PRMT5 antibody.

Figure 3. PRMT5 binds to a single domain in the preLIM region. (A) the diagram shows the progressive deletions of AJUBA. (B) the preLIM region of AJUBA binds to PRMT5. (C) the region between residues 279 and 312 of AJUBA bind to PRMT5. Plasmids were transfected into HEK293 cells. Immunoprecipitations were carried out with anti-Myc antibody, and western blots were performed with anti-Flag antibody.

Figure 4. The leucine residues in the region between residues 279 and 312 of AJUBA are essential for PRMT5 binding. (A) the diagram shows the substitutions for leucine residues in AJUBA. (B) simultaneous mutation of the three leucines to alanine abolishes the interaction between AJUBA and PRMT5. Co-immunoprecipitation was carried out in HEK293 cells as in Figure 2. (C) multiple regions in the PRMT5 protein can bind AJUBA. Bacterially expressed and purified GST-AJUBA (244-350aa) and the indicated

mutant proteins were incubated with *in vitro* translated PRMT5 proteins, and the binding data are summarized in the figure.

Figure 5. AJUBA and PRMT5 are co-localized in U2OS cells. The plasmids encoding Myc-AJUBA and Flag-PRMT5 were transiently transfected into U2OS cells and immunofluorescent images were taken with confocal microscopy. (A) AJUBA is predominately cytoplasmic when expressed alone. (B) the AJUBA L3A mutant is also localized in the cytoplasm and mutation of the three leucines to alanine does not apparently affect its localization. (C) PRMT5 proteins is cytoplasmic with similar distribution patterns to AJUBA. (D) co-expression of AJUBA and PRMT5. (E) co-expression of AJUBA L3A and PRMT5.

Figure 6. The SNAIL, AJUBA and PRMT5 proteins are found in the same complex. (A) the AJUBA protein can simultaneously immunoprecipiate SNAIL and PRMT5. Plasmids encoding Myc-*Ajuba*, Flag-*Snail* and Flag-*PRMT5* were transiently transfected into HEK293 cells, and co-immunoprecipiation was performed with anti-Myc antibody and blotted with anti-Flag antibody. (B) when expressed alone, the SNAIL protein is localized in the nucleus, while the RFP-AJUBA and PRMT5 proteins are in the cytoplasm, the RFP alone is found in both cytoplasm and nucleus. (C) co-expression of RFP and SNAIL does not affect their localization. (D) co-expression of AJUBA and SNAIL affects both of their localization: the SNAIL protein is retained in the cytoplasm, and concomitantly a significant amount of the AJUBA protein is localized in the nucleus. (E) co-expression of PRMT5, RFP and SNAIL reveals no apparent effect on either

localization. (F) co-expression of RFP-AJUBA, SNAIL, and PRMT5 results in a significant amount of the PRMT5 protein being relocated to the nucleus and co-localized with nuclear SNAIL and AJUBA.

Figure 7. Modulation of AJUBA and PRMT5 in p19 cells results in upregulation of *E-cadherin* expression. (A) western blot shows that siRNA targeting AJUBA can stimulate *E-cadherin* gene expression in p19 cells. (B) RT-PCR analysis of the *E-cadherin* mRNA level in si-*Ajuba* p19 cells. (C) the methyltransferase inhibitor 5'-Deoxy-5'-methylthioadenosine (MTA) stimulates *E-cadherin* expression in p19 cells. (D) similarly, siRNA knockdown PRMT5 in p19 cells results in upregulation of *E-cadherin* gene expression shown by western blot.

Figure 8. SNAIL, AJUBA and PRMT5 are associated with the *E-cadherin* gene at the proximal promoter. (A) the diagram ilustrating the *E-cadherin* promoter and PCR primers used for ChIP. A, B and C in the boxes indicates the three SNAIL-binding sites in the promoter. (B) western blot analysis of the Flag-SNAIL, AJUBA, PRMT5 E-cadherin protein expression in HEK293 cells indicates that the *E-cadherin* gene is down-regulated by overexpression of Flag-SNAIL. (C) PCR analysis of the immunoprecipitated DNA fragments.

Figure 9. The model for AJUBA and PRMT5 in SNAIL-mediated gene repression.

Figure 1

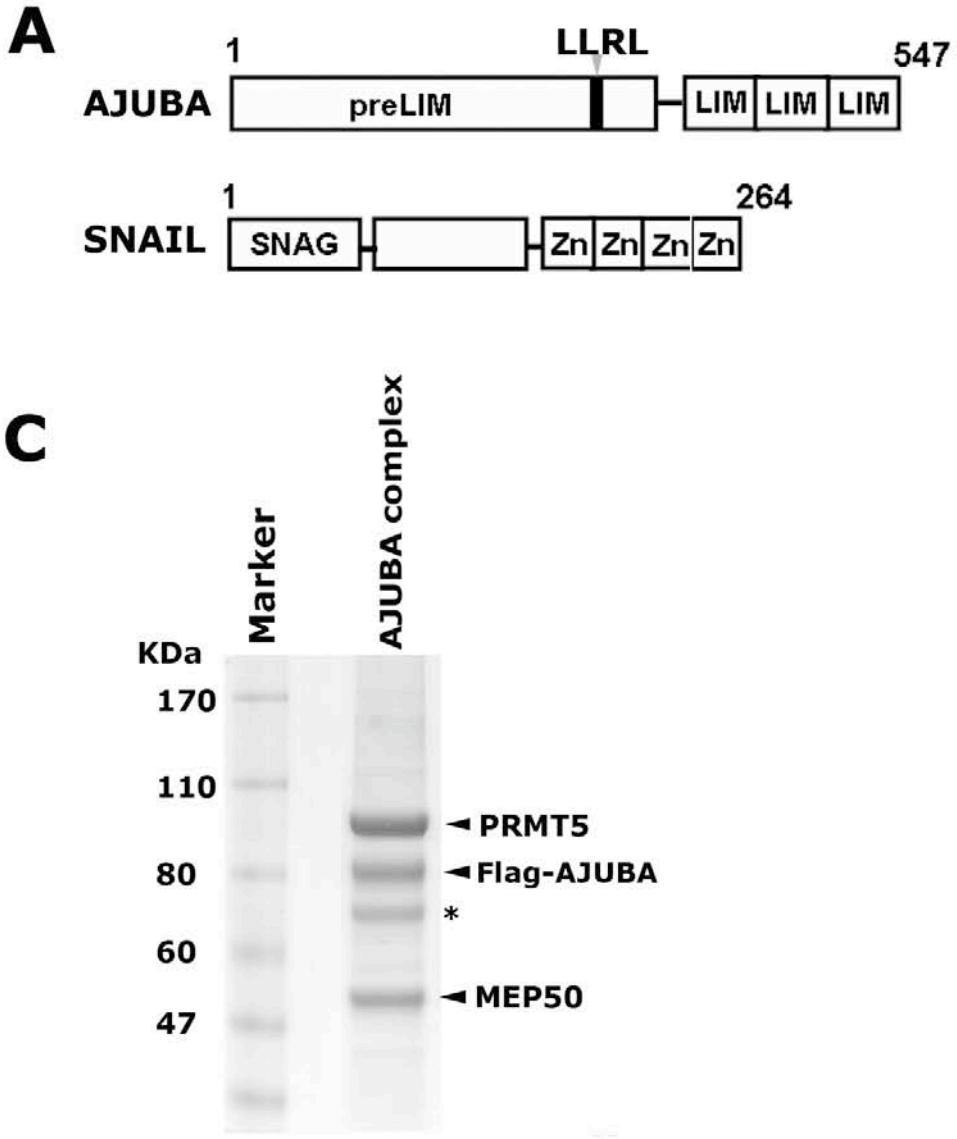

A
1
LLRL
547
AJUBA
preLIM
LIM
LIM
LIM
1
264
SNAIL
SNAG
Zn
Zn
Zn
Zn
C
Marker
AJUBA complex
KDa
170
110
80
60
47
PRMT5
Flag-AJUBA
*
MEP50
B
Luciferase activity (RLU)
35
30
25
20
15
10
5
0
Snail (ng) 0 25 50 100
Luciferase activity (RLU)
25
20
15
10
5
0
Snail (50ng) - + + + +
Ajuba (ng) 0 0 50 100 200

Figure 2

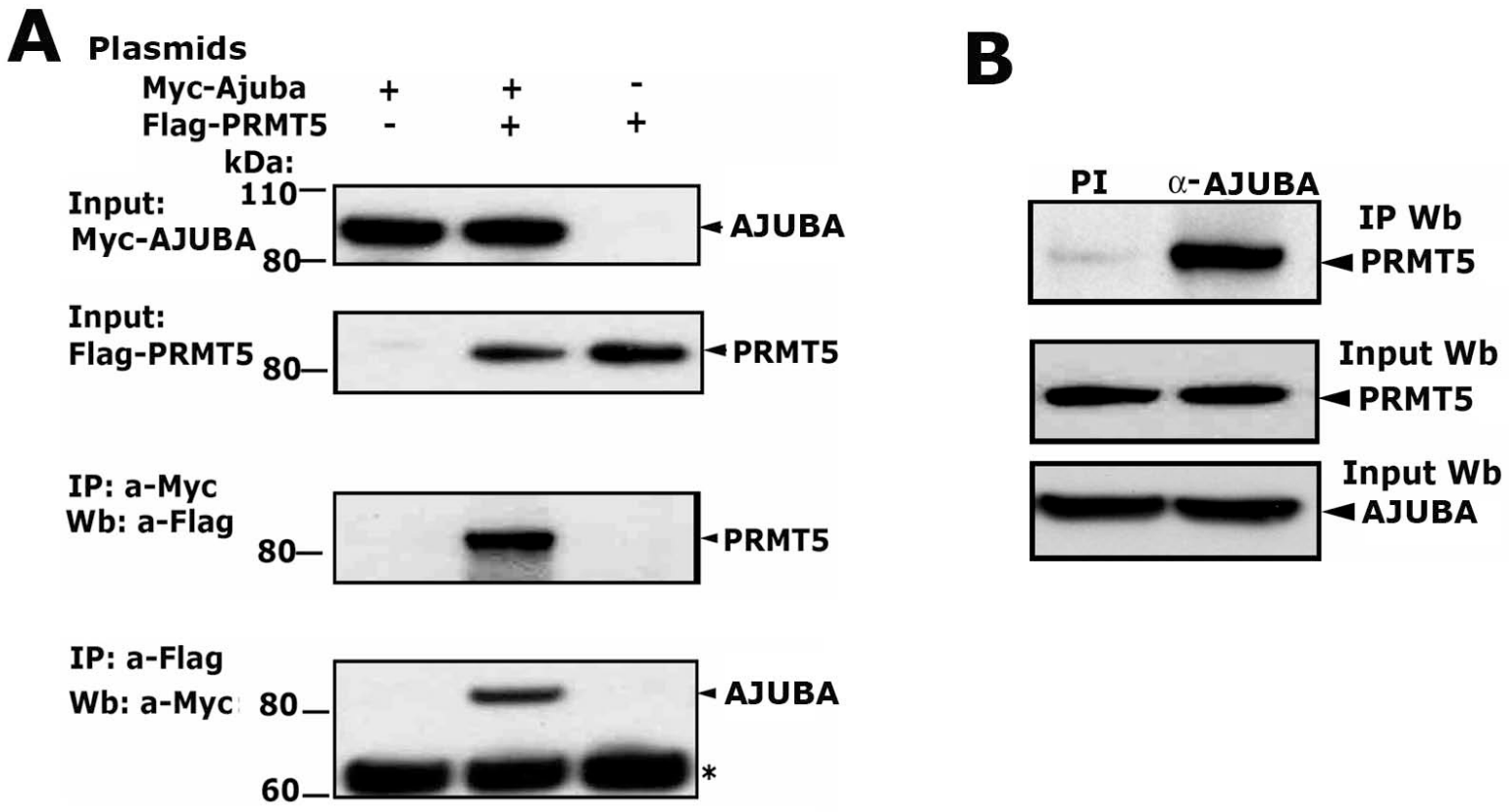

A
Plasmids
Myc-Ajuba + + -
Flag-PRMT5 - + +
kDa:
Input:
Myc-AJUBA
110
80
AJUBA
Input:
Flag-PRMT5
80
PRMT5
IP: a-Myc
Wb: a-Flag
80
PRMT5
IP: a-Flag
Wb: a-Myc
80
60
AJUBA
*
B
PI
α-AJUBA
IP Wb
PRMT5
Input Wb
PRMT5
Input Wb
AJUBA

Figure 3

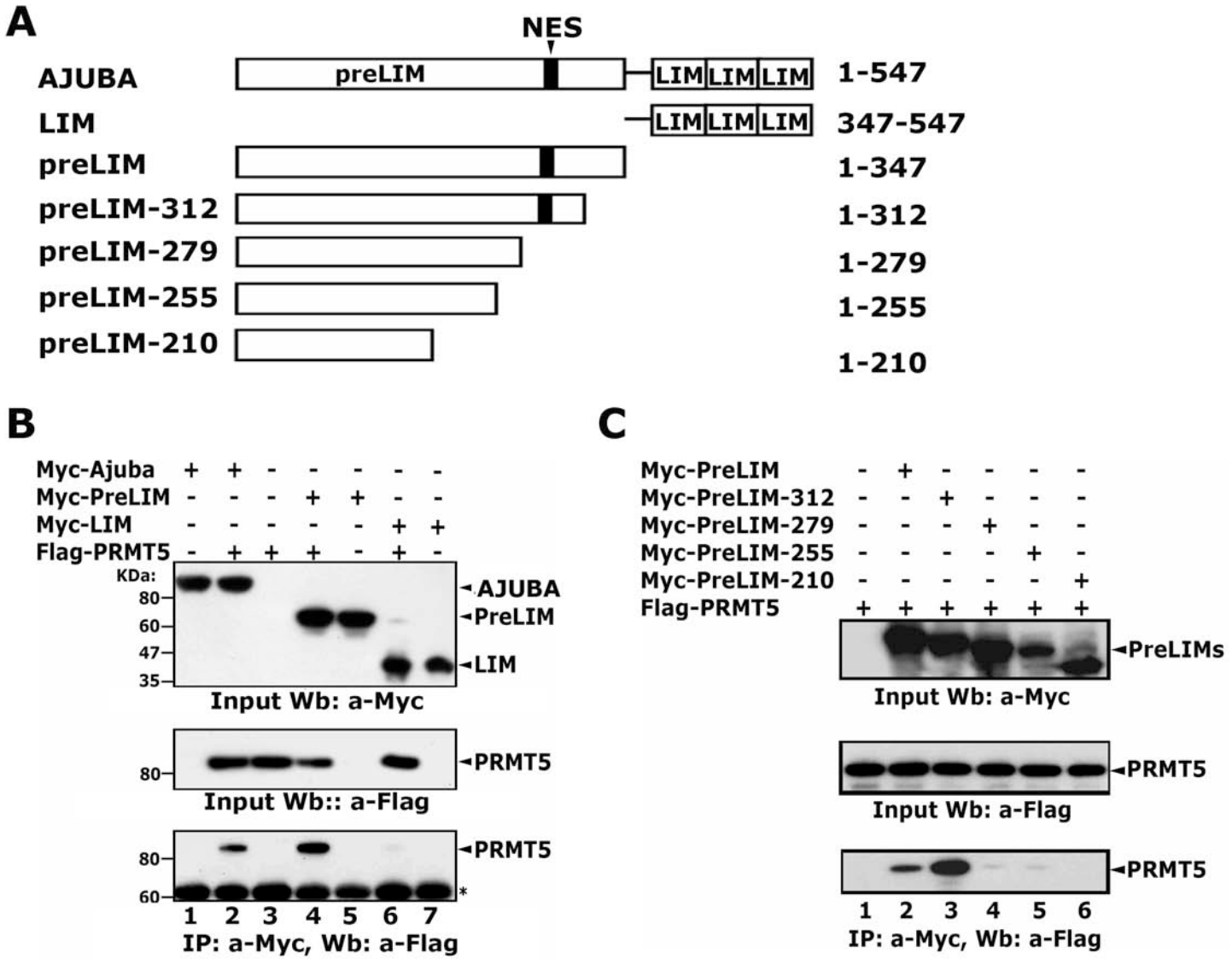

A
NES
AJUBA
preLIM
LIM LIM LIM
1-547
LIM
LIM LIM LIM
347-547
preLIM
1-347
preLIM-312
1-312
preLIM-279
1-279
preLIM-255
1-255
preLIM-210
1-210
B
Myc-Ajuba + + - - - - -
Myc-PreLIM - - - + + - -
Myc-LIM - - - - - + +
Flag-PRMT5 - + + + - + -
KDa:
80
60
47
35
AJUBA
PreLIM
LIM
Input Wb: a-Myc
80
PRMT5
Input Wb:: a-Flag
80
60
PRMT5
*
1 2 3 4 5 6 7
IP: a-Myc, Wb: a-Flag
C
Myc-PreLIM - + - - - -
Myc-PreLIM-312 - - + - - -
Myc-PreLIM-279 - - - + - -
Myc-PreLIM-255 - - - - + -
Myc-PreLIM-210 - - - - - +
Flag-PRMT5 + + + + + +
PreLIMs
Input Wb: a-Myc
PRMT5
Input Wb: a-Flag
PRMT5
1 2 3 4 5 6
IP: a-Myc, Wb: a-Flag

Figure 4

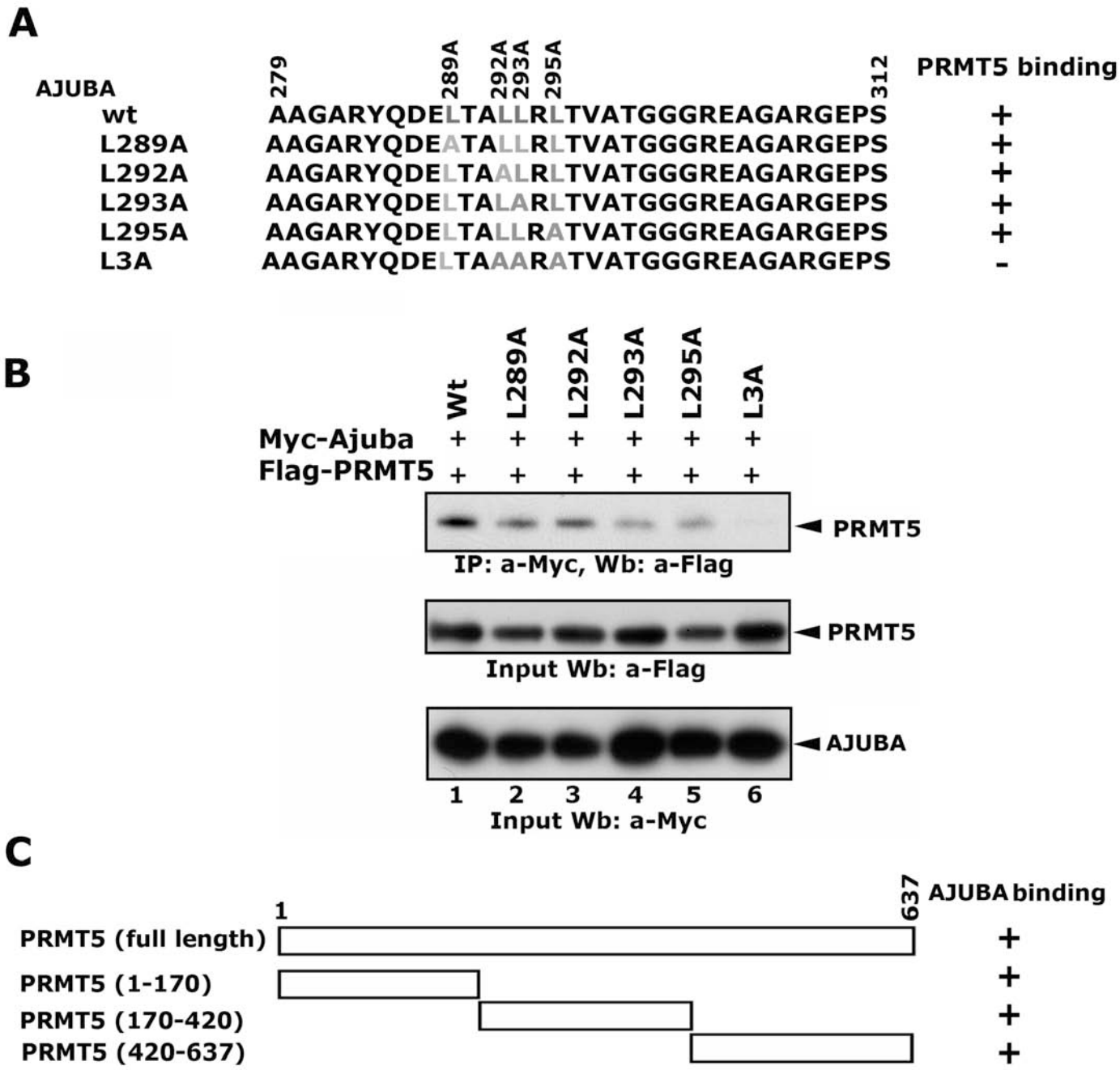

A
AJUBA
279
289A
292A
293A
295A
312
PRMT5 binding
wt AAGARYQDELTALLRLTVATGGGREAGARGEPS +
L289A AAGARYQDEATALLRLTVATGGGREAGARGEPS +
L292A AAGARYQDELTAALRLTVATGGGREAGARGEPS +
L293A AAGARYQDELTALARLTVATGGGREAGARGEPS +
L295A AAGARYQDELTALLRATVATGGGREAGARGEPS +
L3A AAGARYQDELTAAARATVATGGGREAGARGEPS -
B
Wt
L289A
L292A
L293A
L295A
L3A
Myc-Ajuba + + + + + +
Flag-PRMT5 + + + + + +
PRMT5
IP: a-Myc, Wb: a-Flag
PRMT5
Input Wb: a-Flag
AJUBA
1 2 3 4 5 6
Input Wb: a-Myc
C
1
637
AJUBA binding
PRMT5 (full length) +
PRMT5 (1-170) +
PRMT5 (170-420) +
PRMT5 (420-637) +

Figure 5

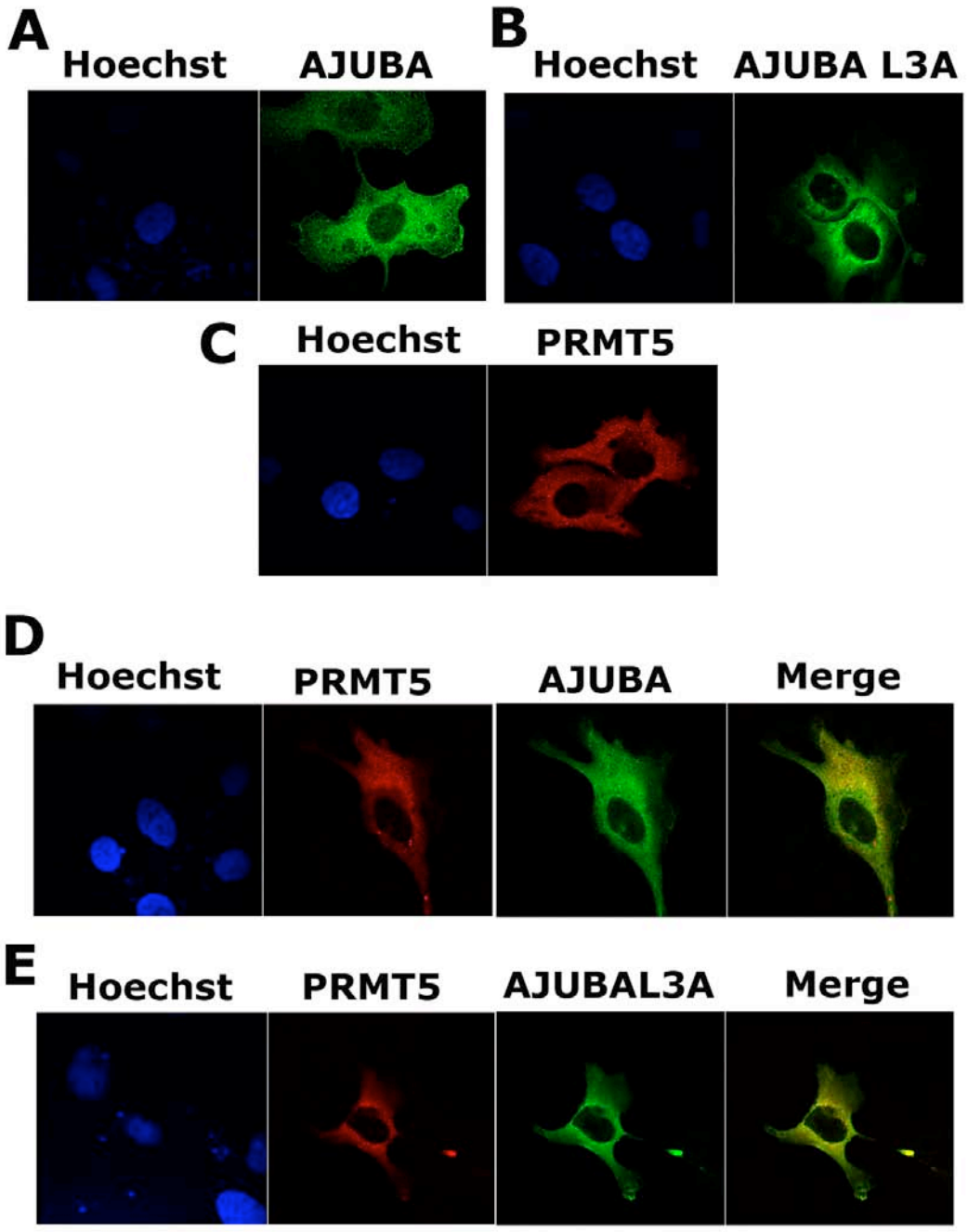

A
Hoechst
AJUBA
B
Hoechst
AJUBA L3A
C
Hoechst
PRMT5
D
Hoechst
PRMT5
AJUBA
Merge
E
Hoechst
PRMT5
AJUBAL3A
Merge

Figure 6

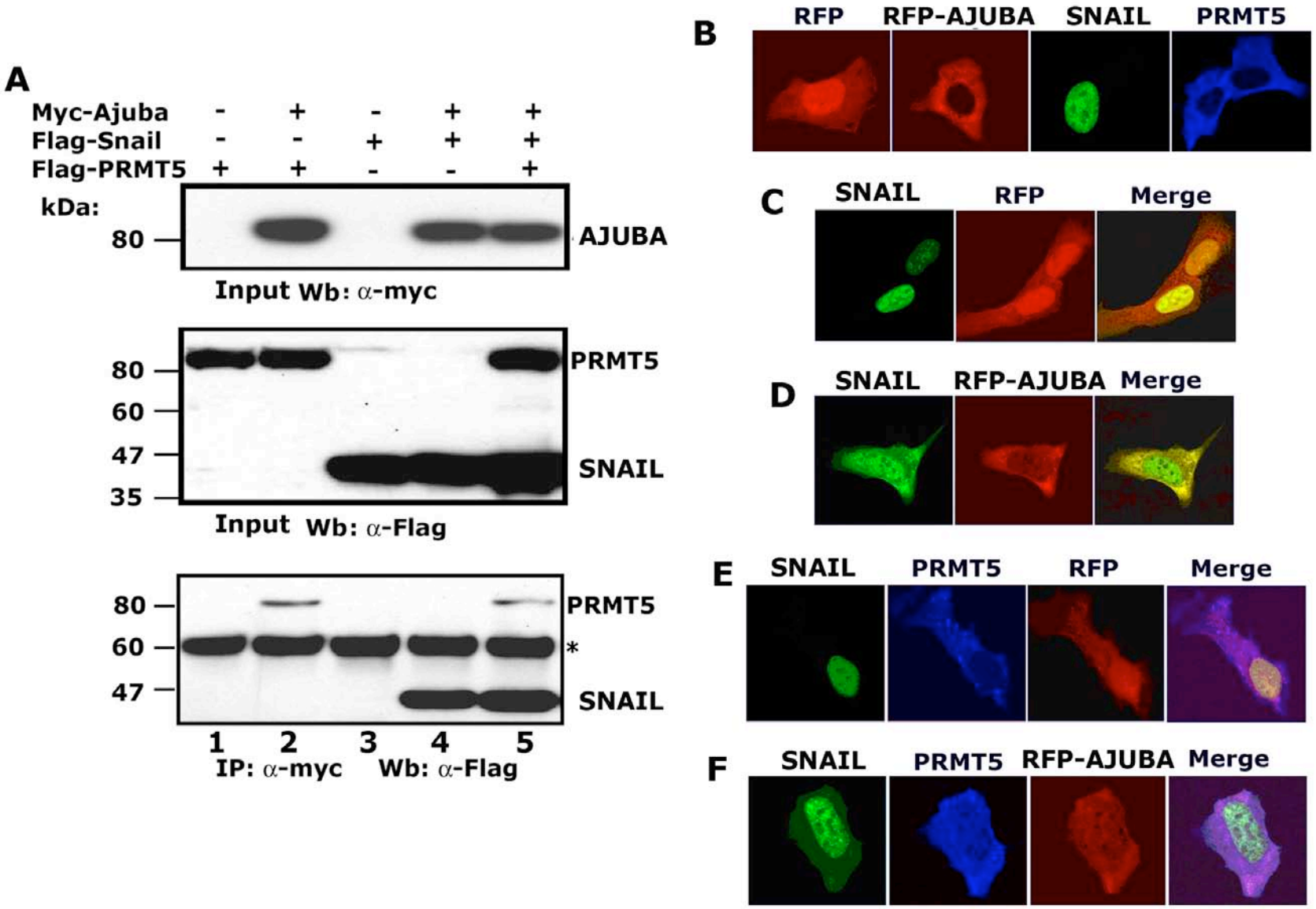

A
Myc-Ajuba - + - + +
Flag-Snail - - + + +
Flag-PRMT5 + + - - +
kDa:
80
AJUBA
Input Wb: α-myc
80
60
47
35
PRMT5
SNAIL
Input Wb: α-Flag
80
60
47
PRMT5
*
SNAIL
1 2 3 4 5
IP: α-myc Wb: α-Flag
B
RFP
RFP-AJUBA
SNAIL
PRMT5
C
SNAIL
RFP
Merge
D
SNAIL
RFP-AJUBA
Merge
E
SNAIL
PRMT5
RFP
Merge
F
SNAIL
PRMT5
RFP-AJUBA
Merge

Figure 7

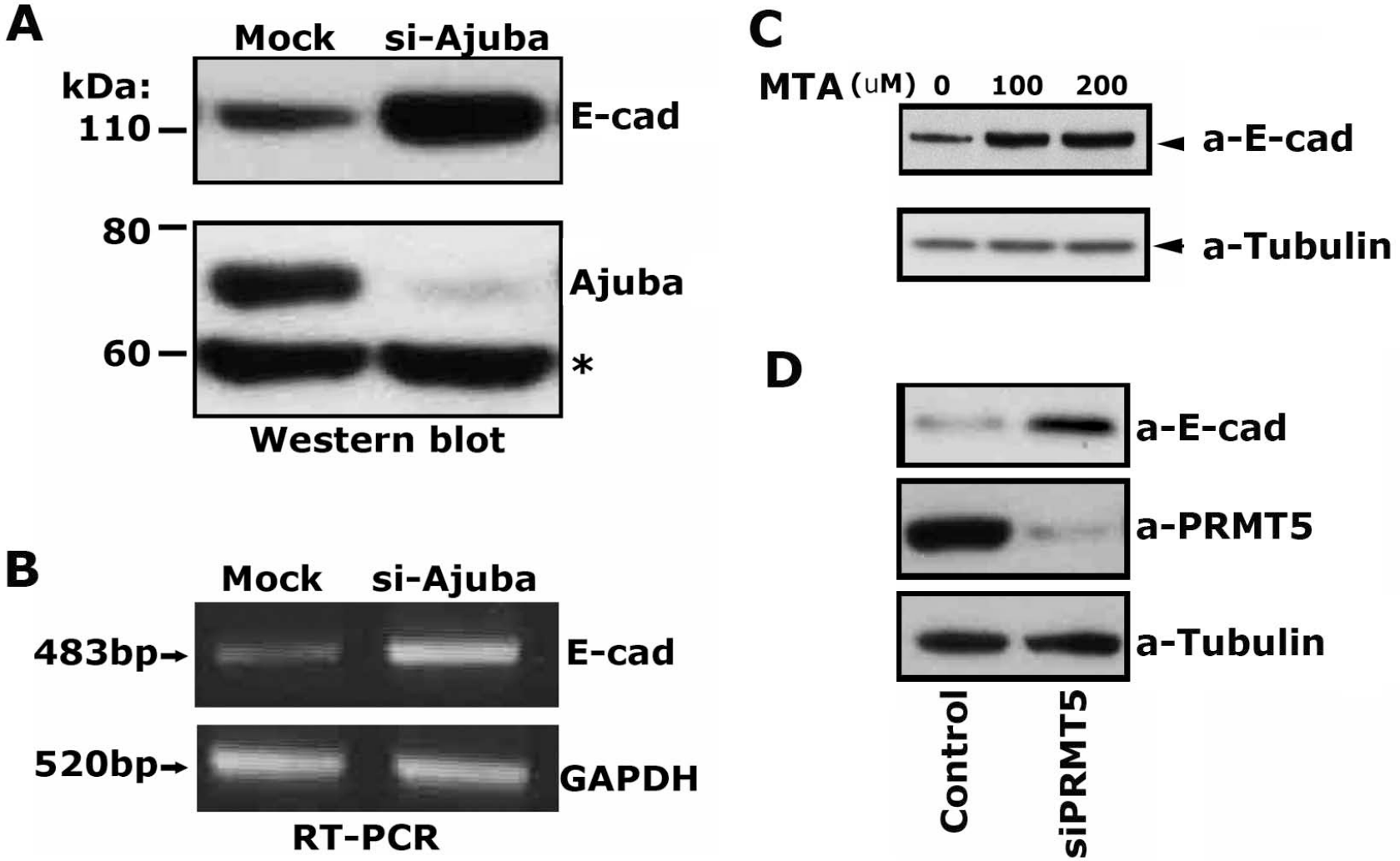

A
Mock
si-Ajuba
kDa:
110
E-cad
80
Ajuba
60
*
Western blot
B
Mock
si-Ajuba
483bp
E-cad
520bp
GAPDH
RT-PCR
C
MTA (uM) 0 100 200
a-E-cad
a-Tubulin
D
a-E-cad
a-PRMT5
a-Tubulin
Control
siPRMT5

Figure 8

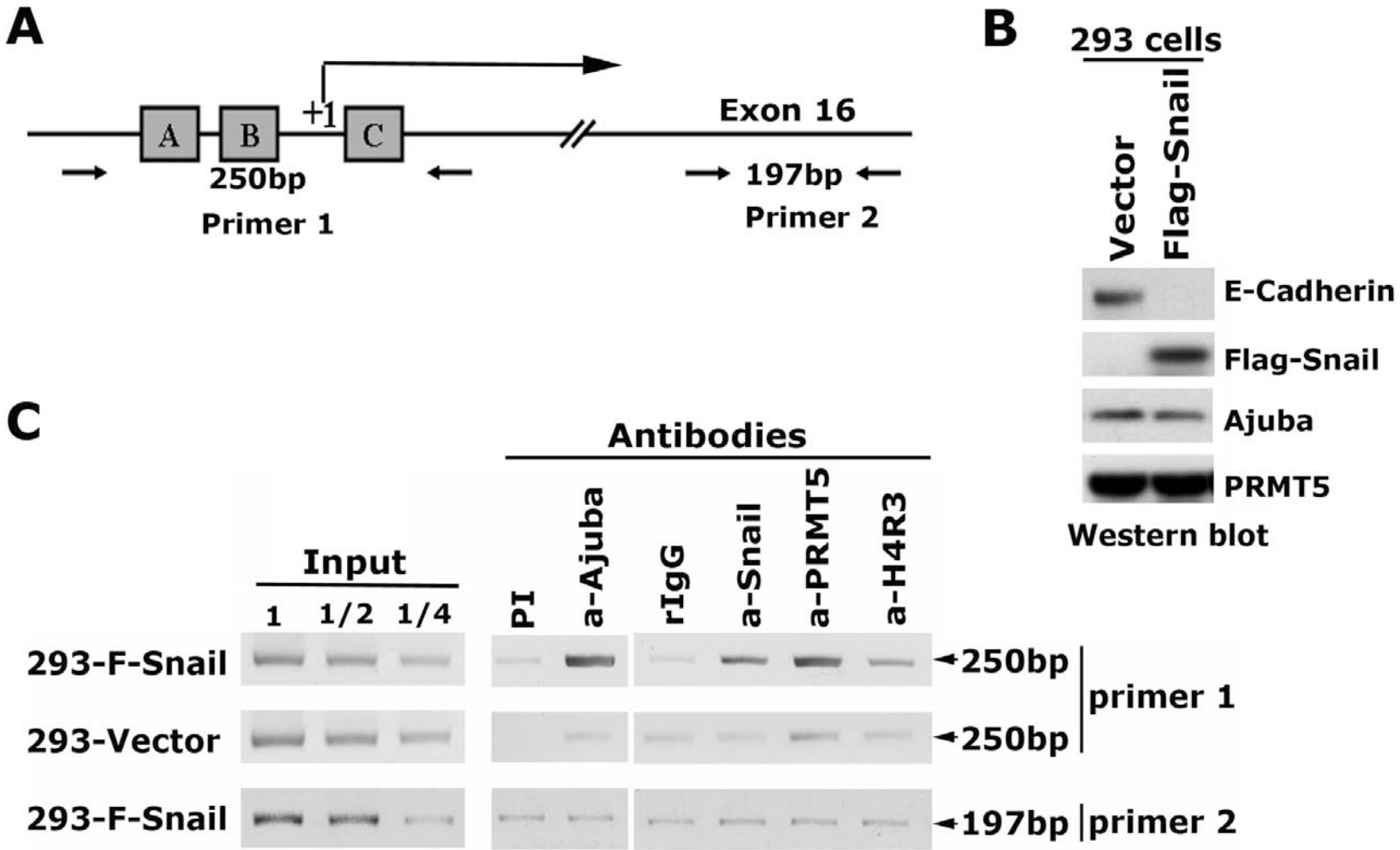
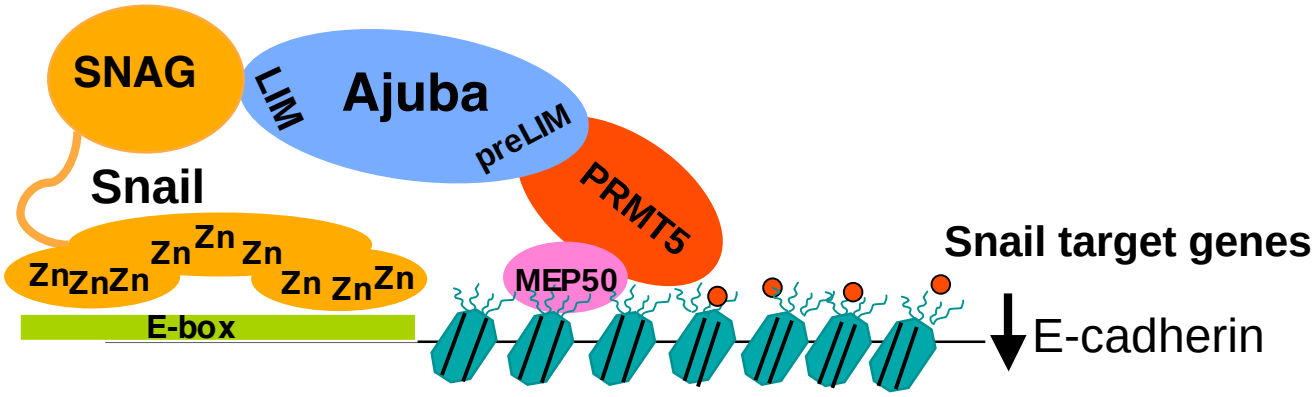

A
A
B
+1
C
Exon 16
250bp
Primer 1
197bp
Primer 2
B
293 cells
Vector
Flag-Snail
E-Cadherin
Flag-Snail
Ajuba
PRMT5
Western blot
C
Antibodies
Input
1
1/2
1/4
PI
a-Ajuba
rIgG
a-Snail
a-PRMT5
a-H4R3
293-F-Snail
293-Vector
293-F-Snail
250bp
250bp
197bp
primer 1
primer 2

Figure 9


SNAG
LIM
Ajuba
preLIM
PRMT5
Snail
Zn Zn Zn Zn Zn Zn Zn Zn
E-box
MEP50
Snail target genes
E-cadherin